



Analyse et Interprétation d'Images : Analyse Multiéchelle pour la Description des Formes Planaires

Kidiyo Kpalma

► To cite this version:

Kidiyo Kpalma. Analyse et Interprétation d'Images : Analyse Multiéchelle pour la Description des Formes Planaires. Traitement du signal et de l'image [eess.SP]. Université Rennes 1, 2009. tel-00446030

HAL Id: tel-00446030

<https://theses.hal.science/tel-00446030>

Submitted on 11 Jan 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITE DE RENNES 1

U.F.R. Structure et propriétés de la matière

Section CNU 61

(Génie Informatique, Automatique et Traitement du Signal)

Rapport de synthèse

présenté pour l'obtention de

l'Habilitation à Diriger des Recherches

par

Kidiyo KPALMA

Titre du mémoire

Analyse et Interprétation d'Images :

Analyse Multiéchelle pour la Description des Formes Planaires

Mémoire présenté le 14 octobre 2009 devant le jury composé de :

Goze Bertin BENIE, Professeur à l'Université de Sherbrooke (Canada),	Rapporteur
Philippe BOLON, Professeur à l'Université de Savoie,	Rapporteur
Jean-Marc BOUCHER, Professeur à l'ENST de Bretagne,	Rapporteur
Véronique COAT, Professeure à l'INSA de Rennes,	Examinatrice
Jean-Louis DILLENSEGER, Maître de conférences, HDR, Université de Rennes 1,	Examineur
Denis HAMAD, Professeur à l'Université du Littoral Côte d'Opale,	Examineur
Joseph RONSIN, Professeur à l'INSA de Rennes,	Examineur

à
mes filles
Morgane,
Razenn,
Gaëlle,
et à leur maman chérie d'amour Sophie que nous aimons de tout notre cœur.
à ma mère,
à la mémoire de mon père
à mes frères et sœurs.

Remerciements

Au terme d'un bilan de mes activités dans l'enseignement supérieur et pendant que je clôture ce document de synthèse, je tiens à exprimer ma gratitude à

Goze Béné, Professeur à l'Université de Sherbrooke (Québec, Canada),

Philippe Bolon, Professeur à l'Université de Savoie,

Jean-Marc Boucher, Professeur à l'ENST de Bretagne et à

Denis Hamad, Professeur à l'Université du Littoral Côte d'Opale

pour m'avoir fait l'honneur d'accepter cette lourde charge d'examiner et de rapporter sur mon dossier.

J'adresse mes sincères remerciements aux collègues

Véronique Coat, Professeure à l'INSA de Rennes, qui a toujours su se montrer agréable tant du point de vue de l'enseignement que de la recherche et avec qui je partage une bonne partie de mes enseignements,

Joseph Ronsin, Professeur à l'INSA de Rennes pour la confiance qu'il m'a toujours témoignée dans les diverses réalisations et qui a toujours apporté un soutien sans faille à l'ensemble des travaux présentés dans ce rapport.

De même que Jean-Louis Dillenseger de l'université de Rennes 1, je les remercie pour avoir accepté d'examiner mon dossier et de participer au jury.

Je remercie toutes celles et tous ceux qui ont contribué à la réalisation des travaux présentés dans ce document : je pense, d'une part, à "mes" doctorants Benoît Ogor, Olivier Thépaut, Omar Diop, Ahsan Ursani et à Mingqiang Yang et, d'autre part, à "mes" stagiaires de PFE ou de Masters Pierre-Yves Tréguier, Mohamed Bakhouch, Alejandro Casado Rodrigues, Antoine Robert, Paul Marguerite, Jean-François Briche, Arnaud Ogier, Borislava Velichkova, Christophe Couturier, Gerry Corrigan, Jeffrey Templeton, Solenn Kervoas, Jari Ryyänen. Sans eux tous ces travaux n'auraient pas eu lieu.

Je n'oublie pas les collègues algériens, le professeur Nasreddine Taleb et Miloud Chikr El Mezouar de l'université de Sidi Bel-Abbès ainsi que Youcef Bentoutou du Centre National des Techniques Spatiales d'Arzew, avec lesquels la collaboration a permis la réalisation de travaux remarquables.

Ce document retrace mes activités de plus d'une quinzaine d'années; durant toute cette période, nombreuses sont les personnes que j'ai côtoyées ou croisées et avec qui j'ai pu collaborer ou tout simplement sympathiser. A ce titre, mes remerciements s'adressent également à tous les collègues de l'INSA, ceux d'hier et ceux d'aujourd'hui, notamment ceux du département EII , car chacun, à sa façon, a contribué à la réalisation de ce travail.

*Pour finir, je vous dis "**Mi gbeekpe ilabalɛ!**" ^(*).*

^(*) "Merci à vous tous" en kabiyè, l' une des 40 "langues" du Togo

Sommaire

Introduction générale.....	11
----------------------------	----

Première partie

Résumé des activités	13
----------------------------	----

I - Curriculum vitæ.....	15
I.1 - Etat-civil.....	15
I.2 - Titres universitaires	15
I.3 - Grades dans l'enseignement supérieur.....	15
I.4 - Spécialités scientifiques	16
I.4.1 - Recherche	16
I.4.2 - Enseignement.....	16
I.5 - Contexte de la recherche	16
I.6 - Prime d'encadrement doctoral.....	17
II - Activités de recherche.....	19
II.1 - Analyse et interprétation des images	19
II.1.1 - Analyse de textures	20
II.1.2 - Traitement des images de télédétection.....	21
II.1.2.1 - Filtrage de speckle, segmentation et recalage	21
II.1.2.2 - Fusion d'images pour la visualisation	22
II.1.2.3 - Applications à la gestion de l'environnement.....	24
II.1.2.3.1 - Détection de la brume sèche dans la zone sahélienne	24
II.1.2.3.2 - Classification d'un couvert végétal dans les images à très haute résolution spatiale.....	25
II.1.3 - Analyse pour la reconnaissance des formes.....	26
II.1.3.1 - Analyse multiéchelle de contours de formes planaires.....	27
II.1.3.2 - Description de formes basée sur l'analyse de la silhouette d'objets.....	27
II.2 - Transfert de technologie.....	28
II.2.1 - Simulation de l'affûtage de scie à bois	28
II.2.2 - Outil pédagogique : TELEDITS ou enseignement des DSP à distance	32
III - Activités d'encadrement.....	35
III.1 - Encadrement de thèses	35
III.2 - Thèses à venir.....	36
III.3 - Encadrement de stages de Masters et PFE.....	37
III.3.1 - Stages de Masters recherche.....	37
III.3.2 - Projet de fin d'études (PFE).....	38

IV - Activités d'enseignement.....	39
V - Responsabilités administratives et scientifiques.....	41
V.1 - Responsabilités scientifiques.....	41
V.2 - Collaborations internationales.....	41
V.3 - Responsabilités diverses	42
V.3.1 - Révision d'articles et animation de sessions.....	42
V.3.2 - Responsabilités administratives et d'intérêt commun	42
VI - Travaux et publications.....	43
VI.1 - Thèse et travaux associés	43
VI.1.1 - Thèse	43
VI.1.2 - Conférences internationales avec comité de lecture [1 article].....	44
VI.1.3 - Conférences nationales avec comité de lecture [2 articles]	44
VI.1.4 - Conférences sans actes [1 Conférence]	44
VI.2 - Publication et travaux réalisés après la thèse.....	44
VI.2.1 - Revues internationales avec comité de lecture [7 articles]	44
VI.2.2 - Revues nationales avec comité de lecture [1 article]	45
VI.2.3 - Livre [1 livre].....	45
VI.2.4 - Chapitres de livres [4 livres]	45
VI.2.5 - Conférences internationales avec comité de lecture [25 articles]	45
VI.2.6 - Conférences nationales avec comité de lecture [6 articles]	48
VI.2.7 - Conférences invitées [1 conférence].....	48
VI.2.8 - Conférences sans acte [1 conférence].....	48
VI.2.9 - Rapports de conventions [7 rapports]	49
VI.2.9.1 - Explosion/Fusion pour la visualisation	49
VI.2.9.2 - Modélisation de dents de scies	49
VI.2.9.3 - Assistance à la Gestion de l'Environnement : projet AGER	49
VI.2.10 - Logiciels [2 logiciels]	49
VI.2.12 - Documents pédagogiques	50
VI.3 - Bilan des travaux dans la période.....	51

Deuxième partie

Rapport de synthèse	53
VII - Analyse multiéchelle de contours de formes planaires.....	55
VII.1 - Introduction	55
VII.2 - Principe de la méthode MSGPR	57
VII.2.1 - Préliminaires	57
VII.2.2 - Structure de la méthode MSGPR.....	57
VII.2.3 - Prétraitement.....	58
VII.2.4 - Séparation des coordonnées	59
VII.2.5 - Détermination du point initial.....	60
VII.2.6 - Lissage de contours	61
VII.2.7 - Ajustement d'échelle	62
VII.2.8 - Définition de la carte CPI	64
VII.3 - Application à la description des formes	66
VII.3.1 - Caractérisation de contours	66
VII.3.2 - La carte CPI et le changement d'échelle.....	67
VII.3.3 - La carte CPI et la rotation	69
VII.3.4 - Application à l'analyse des détails	70
VII.3.5 - La carte CPI et le bruit	71
VII.3.6 - Définition des attributs.....	72
VII.3.7 - Application à la reconnaissance des formes.....	73
VII.3.7.1 - Choix des attributs.....	73
VII.3.7.2 - Mesure de similarité	74
VII.3.7.3 - Exemples d'application à la reconnaissance.....	76
VII.3.7.3.1 - Description du contexte expérimental	76

VII.3.7.3.2 - Forme requête : kk707	79
VII.3.8 - Etude des performances en reconnaissance des formes	80
VII.3.8.1 - Présentation de LAMEFor	80
VII.3.8.2 - Etude comparative avec la méthode CSS.....	81
VII.3.8.2.1 - Pouvoir discriminant	81
VII.3.8.2.2 - Résistance aux fausses alarmes.....	82
VII.3.8.2.3 - Comportement aux déformations élastiques	82
VII.3.8.2.4 - Observations sur les résultats	85
VII.3.8.3 - Application à la lecture automatique de plaques minéralogiques	85
VII.3.8.3.1 - Procédure de reconnaissance de caractères	86
VII.3.8.3.2 - Résultats expérimentaux	87
VII.3.9 - Conclusion	89
VIII - Perspectives.....	91
VIII.1 - En analyse multiéchelle	91
VIII.1.1 - Méthode MSGPR	91
VIII.1.2 - Approche basée sur l'angle tangentiel	92
VIII.1.2.1 - Introduction.....	92
VIII.1.2.2 - Motivations	93
VIII.1.2.3 - Analyse multiéchelle.....	94
VIII.1.2.4 - Définition d'un descripteur : quelques pistes	96
VIII.1.2.4.1 - Solution N° 1 : carte d-TASS sous-échantillonnée	96
VIII.1.2.4.2 - Solution N° 2 : fonction d-TASS sous-échantillonnée	100
VIII.1.2.4.3 - Solution N° 3 : pics de la carte d-TASS.....	100
VIII.1.2.4.4 - Solution N° 4 : distance entre points essentiels par échelle.....	101
VIII.1.2.5 - Conclusion.....	101
VIII.2 - En fusion d'images.....	101
IX - Conclusion générale	103
X - Sigles utilisés.....	105
XI - Bibliographie.....	109
XI.1 - Références bibliographiques	109
XI.2 - Références webographiques	117
XII - Index	119

Troisième partie

Recueil des travaux	121
XIII - Analyse multirésolution d'images de textures	123
XIV - Segmentation d'images de télédétection.....	133
XV - Recalage d'images multirate et multicapteur	143
XVI - Recalage d'images de télédétection.....	153
XVII - Fusion de sur-segmentations	165
XVIII - Descripteur de formes invariant par transformation affine.....	173
XIX - Description de formes par analyse multiéchelle des points d'intersection	181
XX - Description de formes par analyse multiéchelle de l'angle tangentiel	197

Introduction générale

Les activités d'un enseignant-chercheur sont l'enseignement et la recherche. Il contribue ainsi à la création et à la transmission des connaissances. En général, pour le bon fonctionnement de son établissement, ceci se fait conjointement avec des tâches de gestion, des responsabilités administratives et des activités d'intérêt commun. A l'heure de passer une habilitation à diriger les recherches (HDR), étape très importante dans la carrière d'un enseignant-chercheur, l'occasion m'est donnée de faire cette rétrospective sur mes activités.

Issu d'une formation dans la spécialité électronique et informatique industrielle, mes activités d'enseignement et de recherche sont principalement liées aux spécialités du traitement numérique du signal et de l'image. Mes travaux de recherche se placent plus particulièrement et principalement dans le domaine de l'analyse et de l'interprétation des images numériques.

Le présent document comporte trois parties distinctes présentant un aperçu de mes activités d'enseignement et de recherche postérieures à mon doctorat.

La première partie du document comporte six sections. Elle présente un curriculum vitae retraçant ma formation, mes expériences professionnelles et mes domaines de compétences ainsi que le contexte de recherche dans lequel s'est effectué ce travail. Elle donne, également, un résumé de mes charges d'encadrement, d'enseignement et autres types de responsabilités auxquelles j'ai pris part ou dont j'ai encore la charge. Dans la section "Travaux et publications", est présentée une synthèse des publications issues de mes travaux de recherche.

La deuxième partie intitulée "Rapport de synthèse" expose, en détail, les travaux de recherche en cours sur le thème de l'analyse multiéchelle de contours d'objets planaires pour des applications à la description et à la reconnaissance des formes (RdF). Cette partie se termine par une présentation des améliorations potentielles de mon approche ainsi qu'une revue des perspectives à court et à moyen terme de l'évolution de mes travaux dans ce domaine.

Enfin dans la dernière partie, sont présentés huit articles considérés comme représentatifs de nos activités de recherche. Conjointement avec l'étude présentée, ces articles fournissent les informations complémentaires sur l'ensemble de mes activités de recherche.

Première partie

Résumé des activités

I - Curriculum vitæ

I.1 - Etat-civil

Nom..... : *KPALMA*
Prénom : *Kidiyo*
Date et lieu de naissance : *01/01/1962 à Lama-Bou (Togo)*
Nationalité..... : *française*
Sexe..... : *masculin*
Situation familiale : *vit maritalement, 3 enfants*
Adresse professionnelle : *INSA Rennes, 20 av. des buttes de Coësmes, 35043 Rennes*
Téléphone/ Fax..... : *02 23 23 86 59 / 02 23 23 82 62*
Courriel : *kidiyo.kpalma@insa-rennes.fr*
URLs : *http://www.ietr.org et http://www.insa-rennes.fr*

I.2 - Titres universitaires

- Doctorat en *Traitement de l'Information*, spécialité *Traitement d'Images*, soutenu à l'INSA de Rennes en 1992
- DEA en *Traitement du Signal et Télécommunications*, Option *Signal*, à l'Université de Rennes I en 1988
- Maîtrise d'*Electronique, Electrotechnique et Automatique (EEA)* à l'Université de Rennes I en 1987

I.3 - Grades dans l'enseignement supérieur

- Maître de Conférences au département *Electronique Informatique Industrielle (EII)* de l'INSA de Rennes, depuis septembre 1994
- Attaché Temporaire d'Enseignement et de Recherche (ATER) au département EII à l'INSA de Rennes, 1993-94
- Chercheur vacataire à l'INSA de Rennes, 1992
- ATER au département EII à l'INSA de Rennes, 1989-90
- Allocataire d'Enseignement et de Recherche (ALER) au département EII à l'INSA de Rennes, 1988-89

I.4 - Spécialités scientifiques

I.4.1 - Recherche

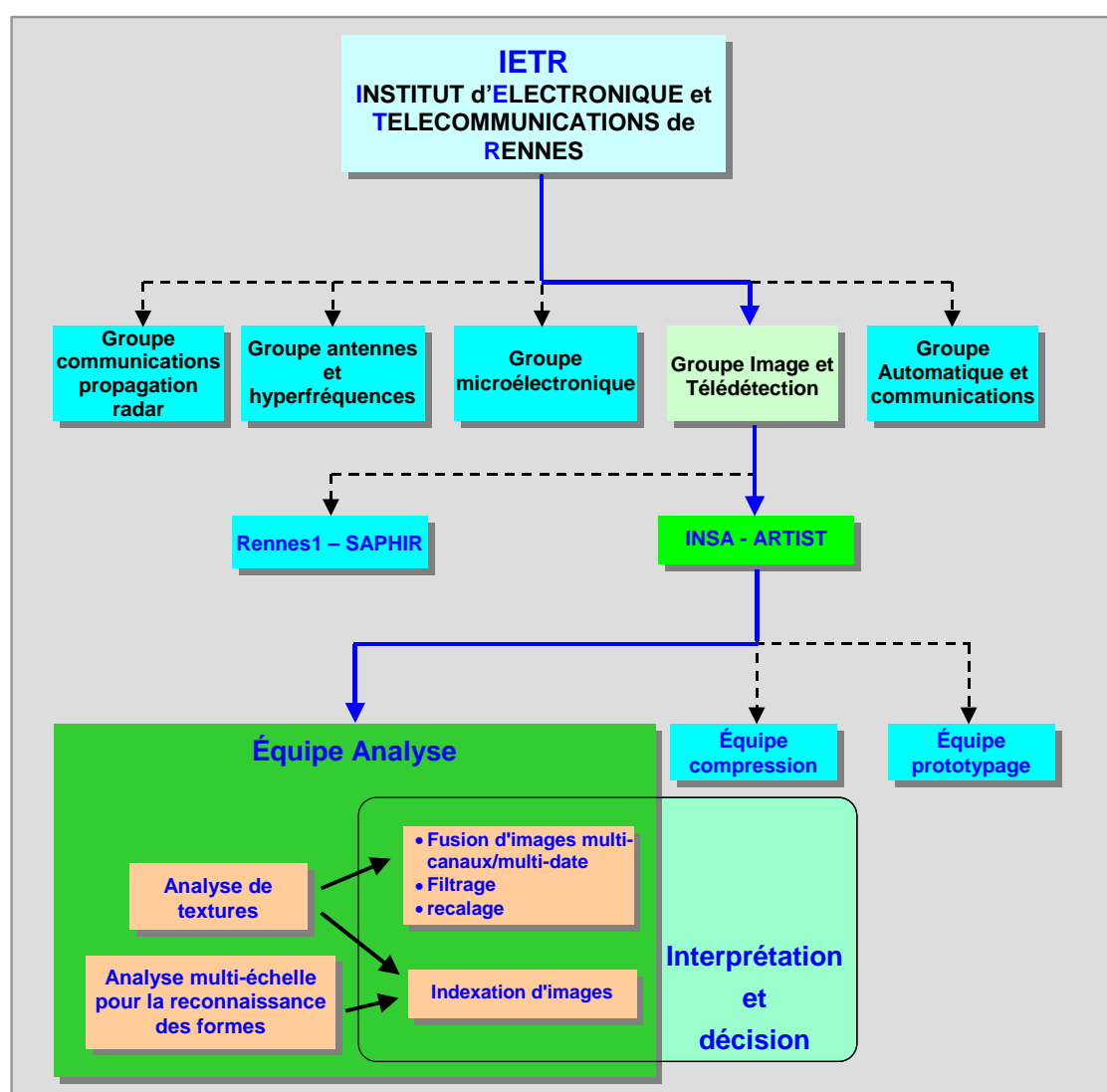
Traitement numérique des images (analyse et interprétation), analyse de textures, segmentation d'images, fusion d'images multicapteurs, multispectrales, multidate, fusion algorithmique, description des formes, reconnaissance des formes.

I.4.2 - Enseignement

Signaux et systèmes, automatique, traitement du signal, processeurs de traitement numérique du signal (DSP), langages de programmation.

I.5 - Contexte de la recherche

Mes activités de recherche se situent dans le groupe *Image et Télédétection* (IT) de l'Institut d'Electronique et de Télécommunications de Rennes (IETR), une UMR associée au CNRS. Cette unité de 240 personnes environ rallie des équipes de recherche en électronique et télécommunications de l'Université de Rennes 1, de l'INSA de Rennes et de SUPELEC, campus de Rennes. L'IETR se compose ainsi de cinq groupes de recherche répartis selon les différents thèmes de recherche, comme le montre l'organigramme ci-dessous.



Organigramme de l'IETR

Comme la plupart des groupes de l'IETR, le groupe *Image et Télédétection* (IT) est réparti sur différents sites dont celui de l'INSA de Rennes auquel j'appartiens. Le groupe IT du site de l'INSA comporte trois équipes travaillant sur trois thématiques différentes qui sont :

- l'équipe ***Compression*** travaille sur les méthodes de représentation basées contenu pour le codage d'images fixes et en séquence,
- l'équipe ***Analyse et interprétation*** s'occupe des méthodes d'analyse et d'interprétation d'images naturelles et d'images de télédétection,
- l'équipe ***Prototypage*** s'intéresse au portage automatique d'algorithmes sur des architectures parallèles et mixtes multi-DSP et FPGA.

Mes domaines de compétences et mes centres d'intérêts sont le traitement du signal en général et le traitement d'images, en particulier. Je m'intéresse plus spécialement à l'analyse des images pour leur interprétation. C'est ainsi que j'ai été désigné pour animer l'équipe Analyse sur la thématique d'*Analyse et interprétation d'images* avec des applications de segmentation d'images, de classification d'images, de reconnaissance d'images...

1.6 - Prime d'encadrement doctoral

Je suis bénéficiaire d'une prime d'encadrement doctoral et de recherche depuis octobre 2008 : ceci indiquant ainsi une reconnaissance de nos travaux de recherche par la communauté scientifique de la spécialité.

II - Activités de recherche

Les travaux de recherche développés dans ce document se placent dans la continuité des études réalisées en analyse de textures pendant ma thèse avec une ouverture vers la reconnaissance des formes. Ces travaux se situent dans l'un des axes de recherche du laboratoire autour de la thématique intitulée *analyse et interprétation d'images* et qui comporte trois thèmes principaux :

- l'analyse de textures,
- le traitement d'images de télédétection,
- l'analyse et description des formes.

Dans ce document, nous aborderons, également, des travaux réalisés en marge de cette thématique : il s'agit, essentiellement, des études réalisées pour répondre à un besoin spécifique, qu'il soit industriel ou universitaire comme la simulation de l'affûtage de scies ou l'outil logiciel pour l'enseignement des DSP à distance.

La section suivante présente un résumé de nos travaux de recherche. Pour simplifier la présentation, ces travaux se déclinent en deux sous-sections : la sous-section "**analyse et interprétation des images**" retrace les travaux directement liés à l'axe principal de recherche et la sous-section "**Transfert de technologie**" qui comporte les autres réalisations.

Cette section est destinée à présenter un résumé descriptif de nos travaux de recherche. Dans la section Analyse multiéchelle de contours de formes planaires, nous reviendrons plus en détail sur l'une des approches développées dans cadre de la description des formes.

II.1 - Analyse et interprétation des images

L'analyse d'images a pour objet de décomposer ("décortiquer") une image afin d'en extraire des informations pertinentes. Les applications sont diverses. En reconnaissance des formes, les informations extraites sont des attributs qui sont sélectionnés pour décrire la forme. La notion de forme, ici, est très vague : peut être un pixel, une région, un objet ou plus généralement, une forme.

II.1.1 - Analyse de textures

On cherche à extraire les attributs caractéristiques (ou signatures) des zones texturées pour des besoins de segmentation, de classification, d'identification... Pendant ma thèse de Doctorat, nous avons réalisé une étude d'analyse d'images dans un contexte multirésolution avec la *dimension fractale* comme signature de la texture [Kpalma92]. L'idée directrice de cette étude était basée sur la définition des fractales qui stipule que les fractales sont des objets extrêmement fragmentés, irréguliers et rugueux dont la dimension fractale est fractionnaire. L'une des principales caractéristiques des fractales est la similitude interne ou, plus généralement, l'*autoaffinité* sur laquelle reposent d'ailleurs la plupart des algorithmes de synthèse d'objets fractals et qui est imagée par cet extrait d'un poème *On Poetry: a Rhapsody* de Jonathan Swift (XVIII^e siècle) :

*"So naturalists observe, a flea
Hath smaller fleas that on him prey,
And these have smaller fleas that bite 'em,
And so proceed ad infinitum."*

A part l'autoaffinité géométrique, il existe l'autoaffinité qui se trouve plutôt dans les propriétés statistiques. C'est le cas du mouvement brownien fractionnaire ou fBm (fractional Brownian motion) [Pentland84]. De par sa définition, le fBm est une fonction isotrope : c'est-à-dire que la dimension fractale reste la même quelle que soit l'orientation d'analyse. Mais grâce aux différentes études expérimentales réalisées sur les textures naturelles, nous avons montré que ces dernières ne sont pas isotropes vis-à-vis du modèle du fBm. Il nous est alors apparu intéressant d'exploiter cette anisotropie dans la définition de la signature de texture [Kpalma91b, Kpalma92]. Les résultats issus de travaux ont permis, par la suite, de définir et de proposer deux nouvelles méthodes de segmentation de textures dans un contexte multirésolution :

- l'Algorithme de Segmentation Multirésolution (ASM) [Kpalma93a] a permis d'atteindre des taux moyens de plus de 98% de bonne classification. Cet algorithme a ainsi montré, grâce à nos différents tests, la supériorité de la multirésolution sur la monorésolution dans le contexte de la segmentation.
- l'Algorithme de Segmentation Hiérarchique (ASH) [Kpalma93b], quant à lui segmente chaque niveau individuellement de façon hiérarchique. En estimant l'indice de confiance de classification d'un pixel, on crée une classe dite de rejet qui est constituée de l'ensemble des pixels dont l'indice de confiance est inférieur à un seuil. Aux niveaux suivants de la pyramide, seuls les pixels correspondant à la classe de rejet sont traités. Ce dernier algorithme a permis de réduire grandement la charge de calcul pour des performances satisfaisantes de l'ordre de 94% de bonne classification.

A la suite de ces travaux, nous avons choisi d'orienter nos recherches vers des applications en télédétection. Nous avons alors intégré le projet AGER (Assistance à la Gestion de l'Environnement par le Radar) qui avait pour but d'apporter une aide pour une meilleure gestion des plans d'épandage. Du point de vue de la technique, ce projet avait pour objectif de fournir des images traitées résultant d'une fusion d'informations provenant de capteurs multiples : LANDSAT ou SPOT pour les images multispectrales, ERS-1 pour les images Radar et le modèle numérique de terrain (MNT).

Ce projet a ainsi marqué notre orientation vers le traitement des images de télédétection avec tous les sujets connexes qu'il comporte comme le filtrage de speckle, le recalage, la fusion, la classification du couvert végétal...

II.1.2 - Traitement des images de télédétection

La gestion et la protection de l'environnement sont deux sujets de plus en plus au cœur des préoccupations mondiales [Ogor97, Berroir98, Hochchild00, Kalluri04]. Dans [Kalluri04], les auteurs présentent un panorama des applications possibles en télédétection à l'aide des images satellitaires : gestion des ressources en eau, gestion des ressources naturelles, gestion et réduction des désastres naturels, gestion de l'agriculture, suivi de l'interaction entre le climat et la santé humaine, suivi du peuplement des zones urbaines, suivi du changement... La constellation de satellites d'ores et déjà présente dans la haute atmosphère est enrichie à chaque nouvelle mise en orbite (SPOT, ERS, Radarsat, Kompsat, Formosat, Cosmos, Ikonos, QuickBird...) avec des résolutions de plus en plus fines. Avec l'avènement des satellites de très haute résolution spatiale, jusqu'à 41cm/pixel en panchromatique pour GeoEye-1, il est désormais possible d'intégrer la notion de textures dans l'analyse d'images de télédétection.

Cette profusion de capteurs génère une quantité phénoménale de données qu'il convient de traiter, que ce soit pour les compresser en vue de leur stockage ou pour leur interprétation.

De par leur contenu spectral, ces données sont soit redondantes soit complémentaires aussi a-t-il été envisagé de combiner les différents canaux dans un processus de fusion [Thépaut00c] ou d'explosion-fusion [Descombes95] afin de produire une information la plus pertinente possible, la plus intelligible possible. Pour illustrer la fusion d'images en télédétection, on peut citer Google Earth et son équivalent français GéoPortail qui sont des logiciels permettant de visualiser une partie de la terre, une ville par exemple, avec une superposition (fusion) de photographies aériennes ou satellitaires, de cartes, du relief...

Pour que ces données puissent être interprétées, il s'avère nécessaire de trouver des méthodes d'analyse automatique ou semi-automatique d'assistance. Dans ce contexte, la fusion d'images de télédétection semble devenir une nécessité compte tenu de l'évolution des techniques de gestion des ressources naturelles et agricoles, ainsi que du grand volume de données à traiter. Malheureusement, les images multidata ou multicapteurs ne sont pas directement superposables, aussi doivent-elles être recalées entre elles. De plus, les images d'un Radar à Ouverture Synthétique (ROS) ou SAR (Synthetic Aperture Radar) qui apportent des informations complémentaires, par tous les temps, aux images du visible comportent le bruit spécifique de *speckle* qu'il convient de réduire par filtrage.

II.1.2.1 - Filtrage de speckle, segmentation et recalage

Dans sa thèse [Ogor97] Benoît Ogor, a mené une étude comparative approfondie de différentes méthodes de filtrage de speckle. Cette étude a permis une mise en œuvre d'un système de segmentation coopérative d'images multicapteurs : visible (capteur SPOT) et radar (capteur ERS).

L'étude a prouvé également que l'association d'un filtre multicanal et d'un filtre spatial, permettait d'obtenir finalement un excellent compromis entre les résolutions radiométrique et spatiale. S'inscrivant dans un schéma plus vaste de fusion de données, les études issues de cette thèse nous ont alors permis de dessiner les contours d'un schéma générique pour les études ultérieures de fusion d'images.

Afin d'apporter une solution au problème de recalage, Olivier Thépaut [Thépaut98a-d, Thépaut99, Thépaut00c] a recensé les approches en recalage d'images afin d'aborder la

fusion d'images. La nature et la complexité des déformations géométriques entre deux images n'étant jamais connues a priori, le choix des points de contrôle et du modèle de déformation est une tâche très délicate qui pèse beaucoup sur la qualité finale du recalage. Le principe général des méthodes utilisant une carte de points est d'identifier, dans les deux images, des points de contrôle qui se correspondent en les appariant. On peut, ensuite, estimer localement les paramètres du modèle de déformation grâce aux positions de ces points.

Pour contourner le problème que constitue la mise en correspondance de données de natures différentes que sont les images SPOT et ERS, Olivier Thépaut a proposé une solution consistant à localiser le pic de corrélation des contours dans les deux images. Pour extraire les données contours, il a opté pour un détecteur adaptatif basé sur le rapport des radiométries moyennes (RRM) [Touzi88a, Touzi88b]. Appliquée à la correction géométrique sur un même référentiel d'une image ERS SAR et d'une image SPOT HRV de la région du Mont-Saint-Michel, cette méthode a donné des résultats très satisfaisants.

Je dois signaler ici, l'abandon de cette thèse qui était pourtant arrivée à terme ; malgré les résultats importants obtenus au cours de cette thèse, le candidat, pour convenance personnelle, a préféré interrompre la rédaction de sa thèse.

Dans ce même contexte, nous avons développé une approche basée sur les points de contrôle en utilisant les invariants et en réalisant le recalage à l'aide de l'algorithme d'interpolation du "thin plate spline" [Bentoutou05].

II.1.2.2 - Fusion d'images pour la visualisation

Cette étude menée dans le cadre d'une convention, signée avec la société IPSIS et le CELAR et regroupant plusieurs unités de recherche, a permis de mettre en application, l'ensemble des expériences acquises en analyse et interprétation d'images. Elle s'inscrit dans la réalisation du poste 2 du projet SIROS/DRI (Détection, Reconnaissance, Identification) qui comptait sept postes. L'objet de cette étude était de recenser et de prospecter les méthodes originales prenant en considération la richesse des données et les méthodes d'explosion-fusion pour proposer une maquette d'aide à l'interprétation des données SAR [Thépaut00b, Thépaut00c, Kpalma01].

Traditionnellement, l'interprète d'images travaille sur l'image SLC (Single Look Complex) en module. Pour mener à bien sa mission, il utilise des images monochromes et s'appuie sur des outils simples de manipulation d'images (LUT, zoom...). Du fait de la grande dynamique des images SAR, l'interprète d'images est, parfois, amené à effectuer un certain nombre de manipulations (visualisation globale-locale) sur les images pour pouvoir en extraire le maximum d'informations.

Pour lui faciliter la tâche et augmenter le "rendement" de l'interprète d'images, il fallait "synthétiser" l'information de manière à améliorer son confort visuel. C'est pour cela que l'algorithme d'amélioration globale/locale a été conçu en intégrant certains concepts liés à la vision humaine.

Le photo-interprète étant un observateur humain, il nous a donc fallu tenir compte des facultés visuelles de l'Homme. Après une large étude des aspects généraux de la vision humaine et de la propriété des images SLC, nous avons développé une méthodologie d'explosion-fusion pour la visualisation. Cette méthodologie consiste à sélectionner, en entrée, les traitements d'explosion répondant au mieux aux besoins opérationnels de l'interprète d'image. Après une combinaison (fusion) des différentes entrées, une projection dans l'espace des propriétés mène à la mise en forme des images destinées à

la visualisation. Pour aider l'interprète d'images dans sa tâche de DRI, une approche intuitive "hiérarchique" est de visualiser l'image globale pour trouver "rapidement" les régions d'intérêt (ROI) contenant une cible, puis de la traiter localement pour en affiner l'analyse. L'algorithme d'analyse locale permet alors de rehausser les objets d'intérêt (points brillants, par exemple) par une table de couleurs adéquate sur un fond de gris.

Grâce au schéma d'explosion-fusion développé, le traitement global a amélioré grandement la lisibilité de l'image à l'aide de la représentation, en couleur, des forts réflecteurs en facilitant :

- le repérage rapide des forts réflecteurs,
- la perception de la radiométrie relative des objets détectés.

De même, on peut dire que l'analyse locale est d'un apport bénéfique par rapport à l'analyse globale, car :

- elle permet de détecter certains forts réflecteurs qui étaient en limite de détection dans l'analyse globale,
- grâce à la représentation sur un fond fortement lissé, elle rend plus aisée, la lisibilité des forts réflecteurs

Sans rentrer dans les détails, la figure II.1 présente un exemple d'analyse locale en vues glissantes (VG). D'une part, elle montre l'apport d'une analyse locale en montrant les détails de la zone d'intérêt par rapport à une analyse globale ; d'autre part, la technique de fusion utilisée a permis de rehausser et de mettre en évidence les points d'intérêt que sont les forts réflecteurs.

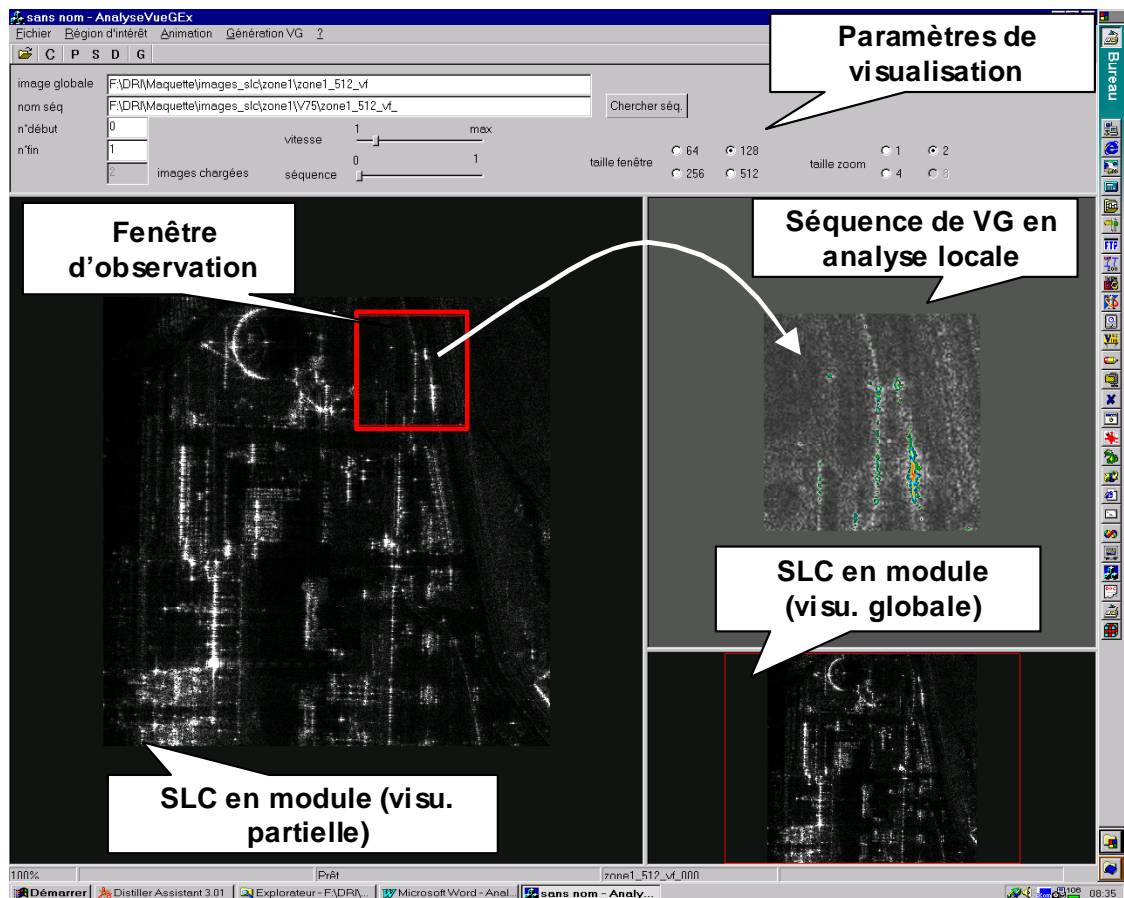


Fig.II.1 : Maquette d'analyse globale-locale

II.1.2.3 - Applications à la gestion de l'environnement

Dans ce cadre, nous avons travaillé sur deux problématiques liées à la gestion de l'environnement : 1) la détection des nuages de poussières ou brume sèche et 2) la classification du couvert végétal sur un site agricole.

II.1.2.3.1 - Détection de la brume sèche dans la zone sahélienne

Dans la zone sahélienne, les activités socio-économiques sont tributaires de la pluviométrie et l'érosion des sols y constitue un problème économique majeur. La quasi-absence des stations de mesure rend difficile la prise en compte de ces phénomènes d'où l'intérêt de trouver de nouveaux outils et de nouvelles voies pour leur détection et la mesure.

L'objectif de Oumar DIOP dans cette étude a été de proposer une solution d'assistance à la gestion de l'environnement dans la zone sahélienne ; nous prenons en entrée le canal infrarouge de Météosat première génération auquel on applique une explosion algorithmique pour ensuite réaliser une fusion qui permet de rehausser et de détecter les nuages de poussières dans les images afin de faciliter leur repérage (Master [Mas2] et thèse [The03]).

La problématique de cette étude est donc la détection des bandes de "nuages de poussières" encore appelés "brume sèche", des positions limites de leurs sources et la configuration de la zone de convergence intertropicale. Ainsi, l'identification des zones limites et des sources des bandes de brume sèche ainsi que la connaissance de la zone de convergence intertropical fournira, entre autres, une aide pour une meilleure gestion des ressources pédologiques [Bertrand77, Jankoviak92] et pour la prévision à court et moyen termes nécessaire aux paysans ou à la sécurité de la navigation aérienne [Dhonneur74, Garba92].

La littérature propose d'autres outils comme l'IDDI [Legrand94] basé sur le canal infrarouge ou le VOAD [N'Tchayi94] extrait du canal visible du satellite Météosat, mais ici, nous proposons d'utiliser une approche associant les attributs de textures dans un schéma d'explosion-fusion [Diop06, Diop06b, Diop07], dans une approche de fusion floue de segmentations. La figure II.2 montre un exemple de résultat obtenu sur un site de la côte ouest de l'Afrique allant du Maroc au Sénégal. En II.2.a, nous avons l'image infrarouge utilisée en entrée et en II.2.b, les résultats de la détection.

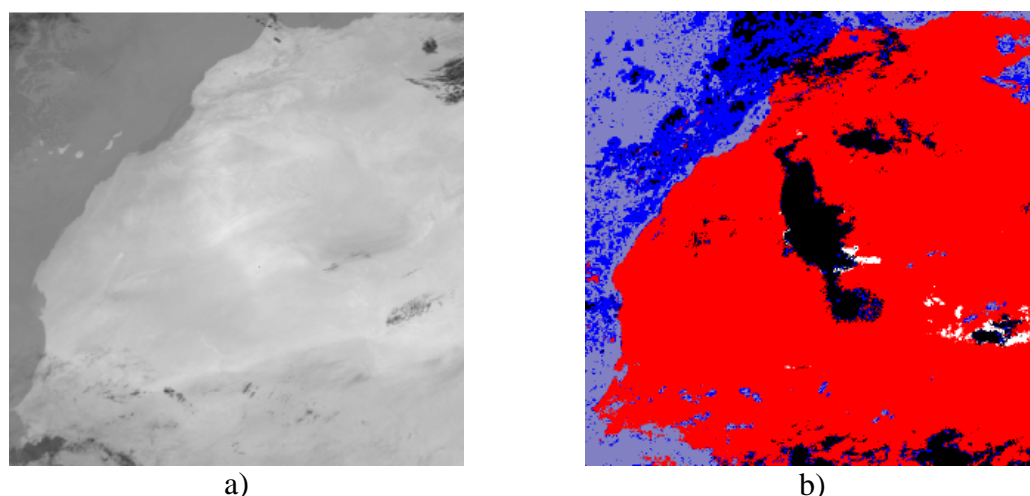


Fig.II.2 : Exemple de détection de nuages de poussière : a) image originale et b) résultat de la détection

Ici, la région d'intérêt est la zone continentale sur laquelle la couleur rouge indique la présence de nuages de poussière tandis que la couleur noir indique l'absence de nuages de poussière. Le mauve indique les nuages d'eau et le blanc la classe d'incertitude définie par l'algorithme de fusion.

II.1.2.3.2 - Classification d'un couvert végétal dans les images à très haute résolution spatiale

Deux contributions apportées en début de thèse de Ahsan URSANI[The01] ont permis de mener à bien le schéma global de fusion : 1) la méthode FOOS (Fusion Of Over-Segmentations) qui apporte une amélioration dans l'exploitation de la méthode des "k-means" [CICL4], 2) apport d'une solution à l'invariance en rotation des caractéristiques texturales issues d'une transformée de Fourier discrète d'une séquence 1D du voisinage 3x3 [RICL1].

Dans cette étude, Ahsan URSANI a mis en œuvre des traitements efficaces en analyse, tant du point de vue spectral et que de la texture, sur des images à très haute résolution (THR) acquises par le satellite Quickbird.

Ce travail combine les résultats de ces deux types d'analyse pour une classification améliorée du couvert végétal. L'analyse spectrale réalisée ici aboutit à une classification non supervisée (segmentation), tandis que l'analyse de texture adopte une procédure de classification supervisée. La fusion des informations de type contour-région, issues de l'analyse spectrale non supervisée, et des informations de type classe, issues de l'analyse texturale supervisée, conduit à des résultats de classification intéressants et encourageants.

Le satellite Quickbird fournit des images dans quatre bandes spectrales : le proche infrarouge (PIR), le rouge (R), le vert (V) et le bleu (B) toutes les quatre de résolution 2.8m/pixel et le panchromatique (PAN) de résolution 70cm/pixel. A partir des trois premières images et de l'image panchromatique, la transformation de Brovey permet d'obtenir une image RVB en fausse couleur de résolution 70cm/pixel. Grâce à ces dernières, des indices de végétation sont calculés. Dans cette étude, et après diverses expérimentations, les indices suivants ont été retenus : l'indice de végétation normalisé (NDVI), l'indice de végétation ajusté au sol (SAVI) et sa version modifiée (MSAVI2) et l'indice normalisé de la différence vert-rouge (NGRDI).

$$NDVI = \frac{\rho_{pir} - \rho_{rouge}}{\rho_{pir} + \rho_{rouge}} \quad (1)$$

$$SAVI = \frac{1.5(\rho_{pir} - \rho_{rouge})}{\rho_{pir} + \rho_{rouge} + 0.5} \quad (2)$$

$$MSAVI2 = \rho_{pir} + 0.5 - 0.5\sqrt{(2\rho_{pir} + 1)^2 - 8(\rho_{pir} - \rho_{rouge})} \quad (3)$$

$$NGRDI = \frac{\rho_{vert} - \rho_{rouge}}{\rho_{vert} + \rho_{rouge}} \quad (4)$$

L'évaluation de l'étude est faite sur des images prises au-dessus d'un site comportant des vergers, des cultures maraîchères, des vignes, des forêts et des jachères de la région de Nîmes (France).

La figure II.3 montre le schéma de fusion utilisé. D'une part, à partir des indices spectraux, FOOS fournit une segmentation en régions homogènes au sens des indices de végétation définis plus haut. Parallèlement, à partir de l'image panchromatique, des

coefficients de Fourier sont extraits d'une séquence 1D du voisinage 3x3 comme expliqué dans [RICL1]. Une découpe en blocs fournit alors les indices texturaux permettant l'identification du contenu végétal de chaque bloc à l'aide de l'algorithme des k plus proches voisins (kNN). Dans l'étape de fusion, un vote majoritaire appliqué à chaque bloc permet de déterminer la classe du contenu d'une région issue de la segmentation.

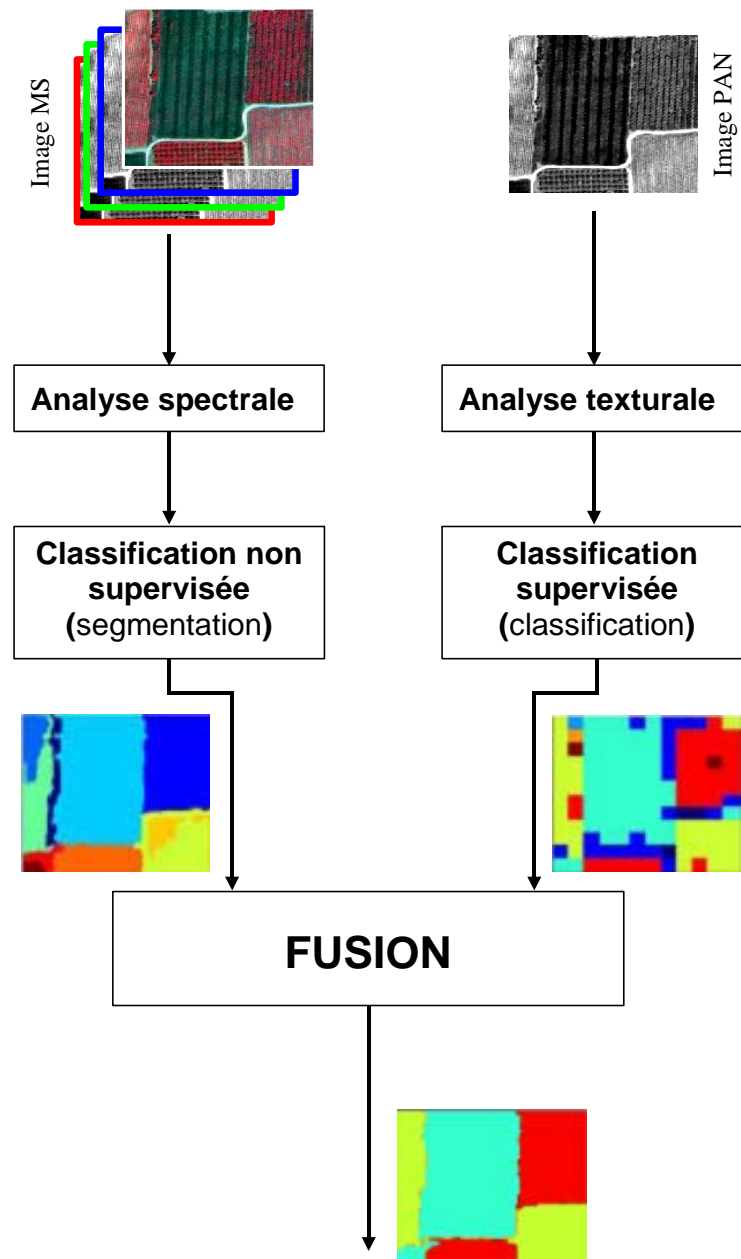


Fig.II.3 : Schéma de fusion de segmentation spectrale et de classification texturale

II.1.3 - Analyse pour la reconnaissance des formes

Dans un monde de plus en plus riche en informations électroniques et où le développement de la technologie conduit à une explosion de l'information, les solutions automatiques de reconnaissance des formes, permettront d'aboutir à des systèmes d'indexation plus efficaces pour les domaines d'application comme : le contrôle qualité de pièces usinées, la détection et la reconnaissance de visages, l'authentification, la

détection, la reconnaissance et le suivi d'objets dans une séquence vidéo, la recherche de formes par le contenu, la lecture automatique des plaques d'immatriculation, la vidéosurveillance...

Dans ce cadre, nous étudions de nouvelles méthodes permettant de caractériser un objet à partir de son contour.

II.1.3.1 - Analyse multiéchelle de contours de formes planaires

Actuellement, nous travaillons sur une méthode originale d'analyse multiéchelle basée sur un lissage progressif du contour par réduction de la largeur de bande du filtre. La détection des points d'intersection entre le contour original et les contours filtrés (après une remise à l'échelle de ces derniers) permet de générer la carte des points d'intersection (CPI) qui est caractéristique de la forme. Les résultats préliminaires encourageants montrent que la carte CPI est une bonne base pour la définition des attributs dans le contexte de la reconnaissance des formes planaires [Kpalma03a, Kpalma03d, Kpalma06a].

Dénommée méthode *MSGPR* (pour *Multi-Scale curve smoothing for Generalised Pattern Recognition*), cette méthode qui est de la famille des méthodes dites d'*espace-échelle* (ou *scale-space*) [Mokhtarian92, Lindeberg94, Matusiak99] permet d'appréhender, à la fois, les caractéristiques spatiales et les caractéristiques d'échelle. Comme dans les autres méthodes de cette famille, le lissage est réalisé à l'aide d'un filtrage passe-bas progressif du contour fermé.

Les résultats expérimentaux obtenus en appliquant cette méthode à divers contours ont montré que cette méthode est invariante à la translation, à la rotation, au changement d'échelle pour une large gamme d'échelle.

Des applications à la récupération d'une forme dans une base de données et une étude comparative avec la méthode CSS ont donné des résultats encourageants et ont montré l'efficacité de notre méthode dans la description et la reconnaissance des formes planaires sans occultation. Dans la deuxième partie du document, nous reviendrons plus en détails sur cette étude.

II.1.3.2 - Description de formes basée sur l'analyse de la silhouette d'objets

Dans ce même contexte, Mingqiang YANG a réalisé une étude ([The02] et [Yang07a, Yang07b]) d'analyse et de description de formes basée tant sur les contours que sur la silhouette des objets. Dans le cadre de cette étude, il a proposé une méthode originale d'échantillonnage d'un contour en se basant non pas sur la distance entre les points consécutifs mais sur l'aire entre deux points consécutifs et le centre de gravité de la forme. En s'appuyant sur ce nouveau type d'échantillonnage, il a développé un descripteur de formes dénommé *Chord Context* et basé sur les statistiques des longueurs de segments (ou cordes) dans la forme dans une direction donnée (voir figure II.4.a). La figure II.4.b montre un exemple de la matrice attribut extraite de la figure II.4.a : dans cette représentation la colonne indique l'angle θ des cordes, la ligne donne la longueur L de la corde et la valeur $M(L, \theta)$, le nombre de cordes de longueur L dans la direction θ . Des tests d'évaluation de ce descripteur, appliqué à la base de données de Kimia (à 99 formes [Sebastian04]) ont donné de très bons résultats comparativement aux descripteurs de référence comme le montre la figure II.4.c (voir [The02]).

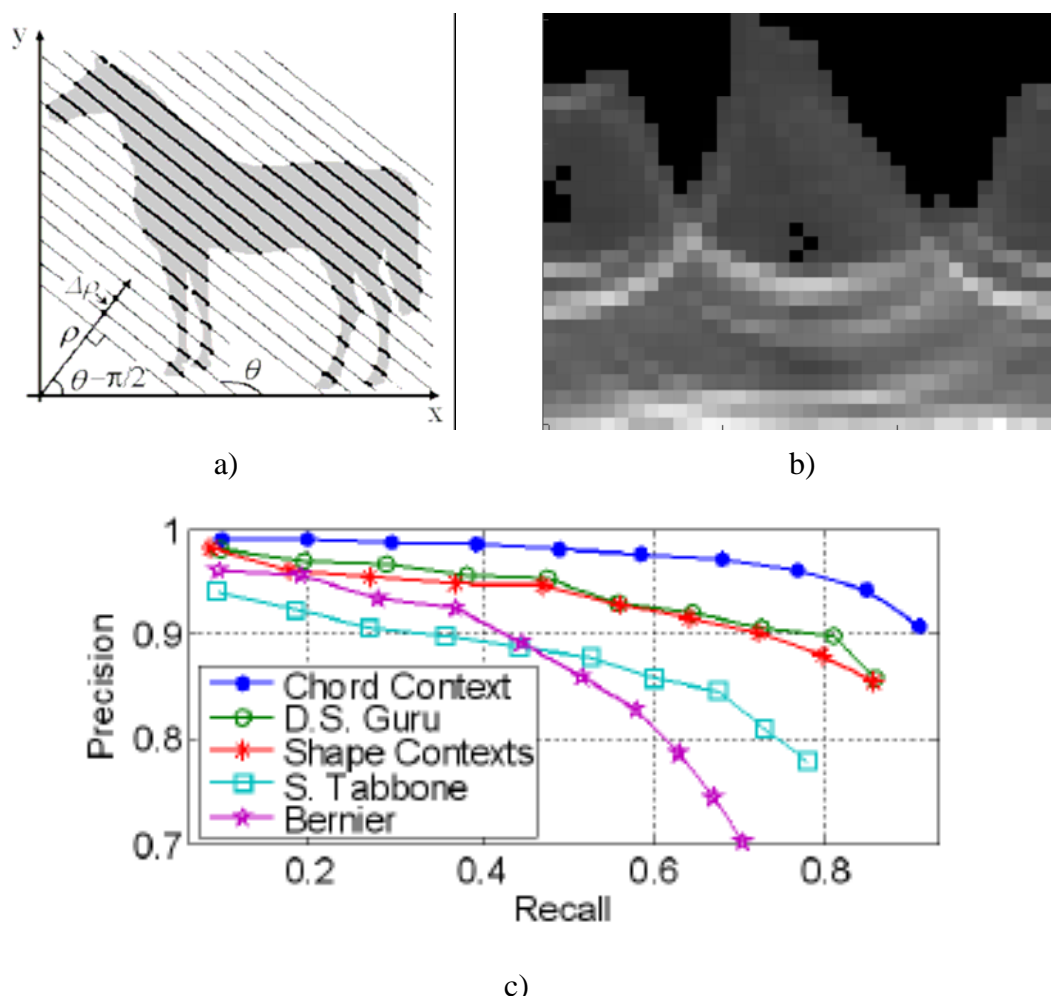


Fig.II.4 : Descripteur Chord context : a) principe et b) matrice attribut et c) graphiques des performances comparativement à d'autres descripteurs existants.

II.2 - Transfert de technologie

Dans cette section, sont présentés les résultats de deux études qui ont été réalisées, l'une dans le cadre d'une convention et l'autre, à titre universitaire dans le but d'aider à l'amélioration de l'enseignement des DSP.

II.2.1 - Simulation de l'affûtage de scie à bois

Cette étude a été réalisée dans le cadre de conventions signées entre 1992 et 1996 avec les Ets Alligator (actuellement, du groupe Vollmer) qui fabriquent et commercialisent des affûteuses de scies à bois. Ce projet a été financé avec le soutien de l'ADEME (Agence Départementale de l'Environnement et de la Maîtrise de l'Energie) [Kpalma92b, Kpalma92c, Kpalma96]. L'objectif de cette étude était de proposer une solution logicielle permettant de générer automatiquement les profils de scies à bois et de meules. Ceux-ci seront, ensuite, intégrés dans un système pour piloter une machine, à commande numérique, dans un processus d'affûtage de scies. Ceci permettant d'alléger les processus d'affûtage, car il faut noter que sur le marché, existent plus de

vingt mille profils de scies différents qui peuvent être circulaires ou à ruban. Dans un système d'affûtage traditionnel, on utilise quatre éléments principaux qui sont :

- la scie (circulaire ou à ruban),
- la meule correspondant au profil de la denture,
- une came pour avancer la lame de scie (déplacement lié au pas de la denture),
- et une came pour déplacer la meule.

De plus, il faut noter que l'avance de la lame est assurée par un poussoir agissant directement sur la dent à affûter. Ceci constitue un inconvénient dans la mesure où cette action directe sur la dent peut détériorer cette dernière. L'inconvénient majeur de l'affûteuse traditionnelle réside dans la multitude (non normalisation) de profils de denture, puisqu'il faut disposer des deux types de comes pour chaque profil. Ceci pose alors le problème de la fabrication de ces comes, de leur stockage et de la gestion de leurs stocks. Pour éviter ces inconvénients, l'idée est donc de fabriquer une affûteuse à commande numérique, car dans ce cas, l'on n'aura plus que quelques paramètres numériques à gérer, et l'on se dispensera complètement des contraintes inhérentes à la gestion et au stockage des comes.

Diverses études nous ont permis de définir dix paramètres pour modéliser le profil d'une scie et quatre paramètres pour décrire complètement une meule [Kpalma92b, Kpalma92c]. Après avoir défini un référentiel de travail pour chacun des deux éléments (scie et meule), nous avons relevé des mesures adéquates sur ces éléments et généré les points de contrôle qui nous ont permis ensuite de synthétiser les profils correspondants. En exploitant les deux profils générés, nous simulons l'affûtage pour en extraire la trajectoire du centre de la meule. La figure II.5, montre en a) la simulation d'un affûtage et en b) une copie d'écran présentant le profil de scie, le profil de meule ainsi que la trajectoire parcourue par le centre de la meule pendant l'affûtage. Les rayons représentés sur le bout de la meule indiquent le taux d'usure de la meule dans la direction correspondante.

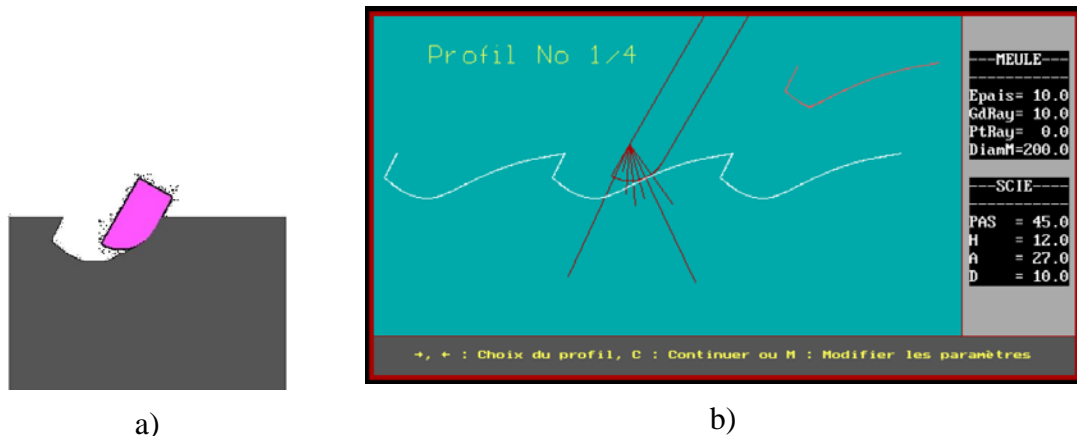


Fig.II.5 : Simulation de la l'affûtage

Sur la figure II.6, est illustré le fonctionnement global du système logiciel intégré à la machine affûteuse. A l'aide des coordonnées de la trajectoire du centre de la meule, le logiciel génère le code pour la machine à commande numérique.

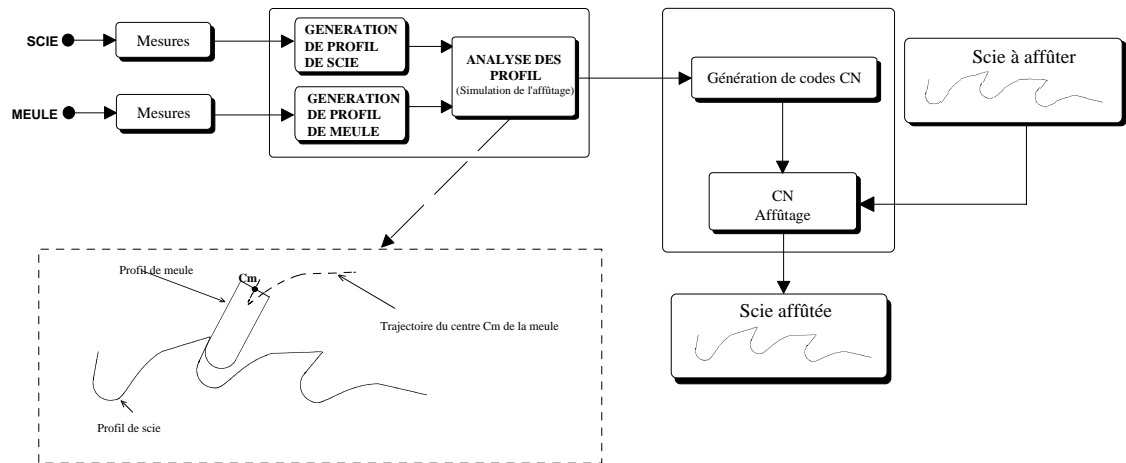


Fig.II.6 : Illustration du processus global de la machine à affûter les scies à bois

Les résultats de cette étude ont permis aux Ets Alligator de fabriquer et de commercialiser une affûteuse à commande numérique.

La figure II.7 suivante montre la copie d'une plaquette présentant la machine à commande numérique qui a été fabriquée à l'aide des résultats issus de cette étude. Les photos des écrans montrent quelques phases dans le processus de simulation et de l'affûtage.

AFFÛTEUSE 2 AXES A COMMANDE NUMERIQUE, POUR LAMES DE SCIE A RUBANS, CIRCULAIRES ET CHASSIS

**Créer et modifier
tous vos profils
en toute liberté ?
Avec Alligator,
vous modélisez,
c'est mémorisé.**



22, AVENUE DE FRIEDLAND
75008 PARIS - FRANCE
TEL.: 33 (1) 42 56 21 78
FAX: 33 (1) 42 25 38 09

- Regardez bien... La nouvelle affûteuse 2 axes à commande numérique d'Alligator est bien plus qu'une évolution majeure, c'est une véritable révolution. En effet, pour la première fois, grâce à son logiciel de modélisation intégré, vous pouvez par simple saisie de 10 paramètres, modéliser le profil exact de la lame à affûter et visualiser ce profil sur écran (ou sur imprimante). Une fois validé, chaque profil modélisé est mémorisé et peut être rappelé très facilement.
- Vous avez donc la totale liberté de modifier vos profils existants ou d'en créer de nouveaux. Pour l'affûtage sous arrosage des lames de scie à rubans, circulaires et châssis, choisir Alligator, c'est plus que jamais anticiper sur le futur.

Création : Nouvion - Photo : Goussier

Fig.II.7 : Machine à commande numérique créée avec les résultats de l'étude

II.2.2 - Outil pédagogique : TELEDITS ou enseignement des DSP à distance

TELEDITS (TELEopération de Dsp par Internet pour le Traitement du Signal)^(*) est un outil qui s'intègre facilement dans un ensemble de *télé-enseignement*. Il vise à mettre du matériel (des DSP) à la disposition des utilisateurs à travers Internet. Ceci permet ainsi aux apprenants de réaliser des travaux pratiques de traitement numérique du signal sur un matériel "physique" distant (e-Labs) [Couturier02, Kpalma03c]. L'outil comporte deux parties interconnectées :

- la partie "**Cours théorique**" est dédiée à la formation théorique sous forme de pages html illustrées par des figures et des simulations en java (<https://ietr-image.insa-rennes.fr/public/Teledits/>).
- la partie "**Téléopération**" est destinée à l'utilisation pratique du matériel en l'occurrence le DSP. Grâce à l'environnement de développement, **DspEdit**, réalisé à cet effet, on peut écrire un programme sur son ordinateur et l'envoyer à un ordinateur distant pour l'y compiler puis le faire exécuter sur le DSP cible. Cet ordinateur distant doit alors être un serveur DSP et héberger le DSP destiné à l'exécution des programmes des clients.

La configuration matérielle du côté d'un fournisseur d'accès à TELEDITS comporte essentiellement des PC dits *serveurs-DSP* qui hébergent, chacun, deux cartes à base de DSP. L'un des DSP est destiné aux traitements requis par le client tandis que l'autre DSP est destiné à jouer le rôle d'un générateur de fonctions téléopérable. Cette configuration donne la possibilité au client de définir les caractéristiques du signal qu'il désire traiter.

Du côté du client, il suffit d'un PC connecté à Internet. Dans les deux cas, les configurations logicielles doivent comporter l'environnement DspEdit.

Outre les fonctionnalités présentées plus haut, DspEdit doit offrir un outil permettant la synthèse de filtres numériques (FltEdit). L'environnement de développement permet également, de générer du code source en langage C, en assembleur ou des scripts Matlab[®] pour des exploitations ultérieures.

Les développements futurs seront essentiellement orientés vers :

- l'amélioration du mode exposé qui consiste à donner la possibilité à plusieurs clients de travailler simultanément sur un même matériel,
- le débogage à distance qui permettra ainsi d'intégrer les notions d'architecture dans TELEDITS,
- l'intégration de l'outil FltEdit de synthèse de filtres dans l'environnement de développement,
- l'élargissement du champ d'application de DspEdit en lui permettant de supporter d'autres types de DSP,
- la refonte[®] du protocole de communication afin de pouvoir se passer de Netmeeting[®] que nous utilisons actuellement pour les communications internet.

^(*) Ce projet a bénéficié du soutien du groupe NTE (Nouvelles Technologies Educatives) de l'INSA de Rennes sous forme de stages de fin d'études et du programme "University Programme" de la société Texas Instruments sous forme de matériel.

Cet outil est actuellement utilisé dans les séances de démonstrations pour illustrer l'utilisation des DSP et pour les travaux pratiques réalisés sur le 'C5X dans le cadre de l'enseignement de la 5^e année au département EII.

La figure II.8.a montre une vue d'ensemble de TELEDITS et illustre les interconnexions possibles en intranet et sur internet. Sur la figure II.8.b est représentée une copie d'écran de l'environnement de développement DspEdit réalisé dans le cadre du projet TELEDITS. On peut y voir les résultats obtenus dans les différentes étapes de la réalisation d'un projet : génération du code source en assembleur, compilation, exécution et récupération des résultats sous forme graphiques.

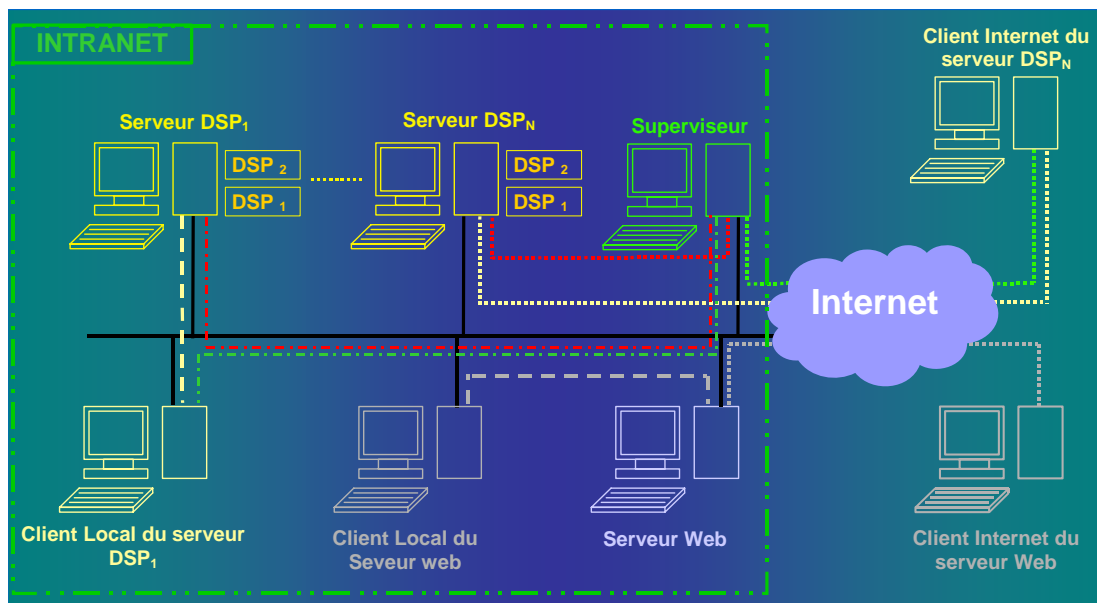


Fig.II.8.a : Vue d'ensemble du projet TELEDITS

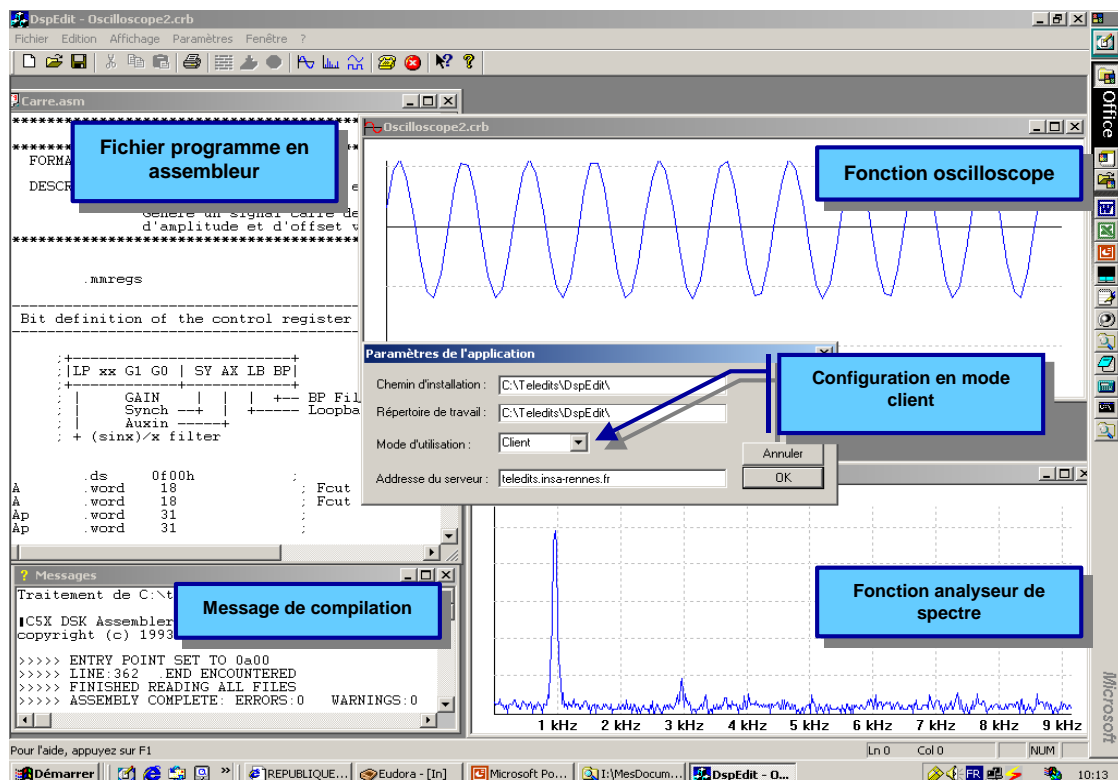


Fig.II.8.b : Environnement de développement de TELEDITS

III - Activités d'encadrement

Cette section présente mes activités en matière d'encadrement qui couvrent les stages de 4^e année (Maters I), de Masters II ou des projets de fin d'études (PFE) d'ingénieurs et bien-sûr l'encadrement des thèses. Ces différents travaux m'ont ainsi permis d'approfondir et d'élargir les principaux aspects de ma thématique de recherche à savoir analyse et interprétation d'images, fusion d'images satellitaires et analyse multiéchelle.

Dans chaque catégorie, chaque élément est présenté sous forme de champs indiquant : le titre de l'étude, l'identité de l'étudiant, les dates de début et de fin de l'étude ainsi que les taux d'encadrement.

Remarque

Pour faciliter la lecture de cette rubrique et la suite du document, nous donnons ici la définition de quelques acronymes qui ne sont pas donnés dans le glossaire :

- **Thex** indique la thèse de rang x dans la catégorie des thèses co-encadrées,
- **Pfex/Masx** indique le projet de fin d'études/Master de rang x dans la liste de cette catégorie des projets encadrés,
- **RICLx/RNCLx** indique l'article de rang x dans la catégorie des articles publiés dans les revues internationales/ nationales,
- **CICLx/CNCLx** indique la communication de rang x dans la catégorie des articles de conférences internationales/nationales avec actes,
- **ChLIx** indique la contribution de rang x dans la catégorie des chapitres de livre,
- **RACox** indique un rapport de convention,
- **COSAx** indique une conférence sans acte,
- **ARSOx** indique un article soumis

III.1 - Encadrement de thèses

[The01] ***Fusion multiniveau pour la classification d'images de télédétection à très haute résolution spatiale***

Nom du diplômé .. : *Ahsan Ursani*

Date début..... : *sept. 2005* Date soutenance: *04 nov. 2008* %encadrements : *70%*

Noms et % des codirecteurs : *Joseph Ronsin (30%)*

[The02] ***Extraction d'attributs et mesures de similarité basées sur la forme***

Nom du diplômé .. : *Mingqiang Yang*

Date début..... : *sept. 2005* Date soutenance : *03 juil. 2008* %encadrements : *70%*

Noms et % des codirecteurs : *Joseph Ronsin (30%)*

[The03] ***Fusion et interprétation d'images satellitaires pour la gestion de l'environnement dans la zone sahélienne***

Nom du diplômé .. : *Oumar Diop*

Date de début : *sept. 2002* Date soutenance : *14 déc. 2007* %encadrement : *70%*

Noms et % des codirecteurs : *Joseph Ronsin (30%)*

[The04]* ***Fusion de données satellitaires ERS, SPOT et cartographiques IGN : application au recalage géodésique, à la segmentation coopérative et à la classification consensuelle d'images multi-sources de la zone des polders du Mont-Saint-Michel***

Nom du diplômé .. : *Olivier Thépaut*

Date de début .. : *sept. 1996* Date fin : *interrompue en 2000* %encadrement : *70%*

Noms et % des codirecteurs : *Joseph Ronsin (20%)*

[The05] ***Etude comparative de méthodes de filtrage de speckle et de segmentation en imagerie radar à synthèse d'ouverture : application à la région des polders du Mont Saint-Michel***

Nom du diplômé .. : *Benoît Ogor*

Date de début : *sept. 1992* Date fin : *déc. 1997* % Encadrement : *50%*

Noms et % des codirecteurs : *V. Haese-Coat (30%) et J. Ronsin (20%)*

III.2 - Thèses à venir

Dès la rentrée 2009/2010, nous allons accueillir deux doctorants sur les sujets suivants :

[The06] ***Fusion d'images en télédétection satellitaire*** avec une bourse franco-algérienne.

Etudier et développer de techniques et algorithmes de fusion efficace d'images satellitaires pour aboutir à des images à très haute résolution contenant le maximum d'informations utiles sous les contraintes spatiales et spectrales fixées par les différents types d'images.

[The07] ***Analyse d'images compressées pour l'interprétation et l'indexation de leur contenu*** : bourse chinoise (CSC-INSA-UT)

En s'appuyant sur le LAR (Local Adaptive Resolution), étudier et proposer un descripteur sémantique. Le LAR fournit une représentation sémantique en régions avec des descripteurs visuels (contours, texture, couleur). L'objectif est donc d'ajouter des fonctionnalités d'indexation en se basant sur des descripteurs de régions qui seront extraits des données compressées.

* Cette thèse n'a pas été soutenue, car le candidat a, volontairement et pour des convenances personnelles, abandonné la rédaction de sa thèse pour un emploi salarié en entreprise.

III.3 - Encadrement de stages de Masters et PFE

Cette section présente les deux types principaux de stage que j'ai encadrés à savoir les projets de fin d'études (PFE) ou les stages de Masters.

III.3.1 - Stages de Masters recherche

- [Mas01] ***Intégration de détection et suivi de contours dans la plate-forme LAMeFor***
 Nom du diplômé : *Hoang Anh NGUYEN*
 Diplôme : *Master Recherche et PFE*
 Date de début : *fév. 2007* Date fin : *juin 2007* % Encadrement .. : *100%*
- [Mas02] ***Traitement d'images radar pour la détection de cibles***
 Nom du diplômé : *Pierre-Yves Tréguier*
 Diplôme : *Master Recherche*
 Date de début : *mars 2005* Date fin : *juin 2005* % Encadrement : *50%*
 Noms et % des codirecteurs : *Stéphane Méric (50%)*
- [Mas03] ***Développement d'outils d'analyse d'images de satellites météo pour la détection des nuages de poussière***
 Nom du diplômé : *Mohamed Bakhouché*
 Diplôme : *Master recherche*
 Date de début : *mars 2005* Date fin : *juil. 2005* % Encadrement .. : *100%*
- [Mas04] ***Développement d'un algorithme d'analyse de courbures multi-échelles pour la reconnaissance des formes***
 Nom du diplômé .. : *Antoine Robert*
 Diplôme : *DEA*
 Date de début : *mars 2004* Date fin : *juil. 2004* %Encadrement ... : *100%*
- [Mas05] ***Compression d'image SAR***
 Nom du diplômé : *Jean-François Briche*
 Diplôme : *DEA et PFE*
 Date de début : *fév. 2003* Date fin : *juin 2003* %Encadrement : *50%*
 Noms et % des codirecteurs : *Olivier Déforges (50%)*
- [Mas06] ***Fusion coopérative d'images satellitaires radar-optique***
 Nom du diplômé .. : *Arnaud Ogier*
 Diplôme : *DEA*
 Date de début : *fév. 2002* Date fin : *juin 2002* %Encadrement ... : *100%*
- [Mas07] ***Etude comparative d'algorithmes de segmentation en imagerie SAR : application à l'extraction de parcelles agricoles***
 Nom du diplômé .. : *Olivier Thépaut*
 Diplôme : *DEA*
 Date de début : *mars 1996* Date fin : *sept. 1996* %Encadrement ... : *100%*
- [Mas08] ***Extraction de structures en imagerie satellitaire et radar en vue de la fusion d'images***
 Nom du diplômé : *Solenn Kervoas*
 Diplôme : *stage de DEA et PFE*
 Date de début : *mars 1995* Date fin : *sept. 1995* %Encadrement ... : *100%*

III.3.2 - Projet de fin d'études (PFE)

- [Pfe01] **Développement d'une Plate-forme Logicielle d'Analyse Multi-Échelle pour la Reconnaissance des Formes (LAMEFor)**
 Nom du diplômé .. : *Alejandro Casado Rodrigues*
 Diplôme : *PFE-ERASMUS*
 Date début : *oct. 2004* Date fin : *avril 2005* % encadrements .. : *100%*
- [Pfe02] **Analyse multi-échelle de formes planaires pour leur reconnaissance**
 Nom du diplômé .. : *Paul Marguerite*
 Diplôme : *PFE*
 Date de début : *mars 2004* Date fin : *sept. 2004* %Encadrement ... : *100%*
- [Pfe03] **Analyse et Reconnaissance des Formes d'Objets en Vrac Planaires et en 3D**
 Nom du diplômé .. : *Elitza Enguelova*
 Diplôme : *Stage post-ingénieur*
 Date de début : *déc. 2001* Date fin : *janv. 2003* %Encadrement : *40%*
 Noms et % des codirecteurs : *Joseph Ronsin (INSA à 30%) et Pencho Venkov de (Université Technique de Sofia en Bulgarie à 30%)*
- [Pfes04] **TELEDITS : Intégration d'un outil de synthèse de filtres dans le logiciel DspEdit**
 Nom du diplômé .. : *Borislava Velichkova*
 Diplôme : *PFE (UTC de Sofia (Bulgarie))*
 Date de début : *mars 2002* Date fin : *juin 2002* %Encadrement ... : *100%*
- [Pfe05] **Téléopération des DSP par Internet – Mode client**
 Nom du diplômé .. : *Christophe Couturier*
 Diplôme : *PFE*
 Date de début : *fév. 2001* Date fin : *juin 2001* %Encadrement ... : *100%*
- [Pfe06] **Internet Based Digital Signal Processing Tutorial**
 Nom du diplômé .. : *Gerry Corrigan*
 Diplôme : *PFE (ERASMUS)*
 Date de début : *sept. 1999* Date fin : *déc. 1999* %Encadrement ... : *100%*
- [Pfe07] **Development of a Graphical User Interface (GUI) for TMS320 DSP chip control**
 Nom du diplômé .. : *Jeffrey Templeton*
 Diplôme : *PFE (ERASMUS)*
 Date de début : *sept 1999* Date fin : *déc. 1999* %Encadrement ... : *100%*
- [Pfe08] **Local Histograms for Texture Characterization and Classification**
 Nom du diplômé .. : *Jari Ryyänen*
 Diplôme : *PFE (Master's Thesis INSA and University of Oulu, Finland)*
 Date de début : *mars 1992* Date fin : *nov. 1992* %Encadrement : ... *100%*

IV - Activités d'enseignement

L'ensemble de mes enseignements est centré sur les disciplines de l'électronique et de l'informatique industrielle : génie "*informatique, automatique et traitement du signal*" et de l'*électronique* à l'INSA de Rennes. Commencées en période doctorale en qualité d'ALER puis d'ATER, ces activités remontent à septembre 1988 avec ma participation à l'encadrement des travaux pratiques de diverses matières puis elles ont évolué progressivement vers des travaux dirigés et des cours magistraux avec la mise en place de nouveaux enseignements, lorsque cela est nécessaire.

Sans chercher à être exhaustif, les paragraphes suivants donnent un aperçu rapide des matières dans lesquelles j'interviens.

Signaux et systèmes : dans l'esprit d'uniformité de l'enseignement que je dispense, j'ai pris en charge le cours de *Signaux et Systèmes*, depuis la rentrée 1998/99. Cet ensemble de cours (32hTD) et TD (30hTD) dispensé aux étudiants de 3^e année, du département EII, leur fournit les outils mathématiques nécessaires pour aborder les cours de *Traitement du Signal* et d'*Automatique*.

Traitement du signal : depuis la rentrée 1995/96, je partage la responsabilité du cours de *Théorie et traitement du signal* en 3^e année et de *Traitement du Signal Numérique* en 4^e année du département EII qui comporte des cours (32hTD), des TD (30hTD) et des TP (18hTD).

Automatique : je participe au cours d'*Automatique* en 3^e année par le biais des TD (18hTD) et des TP (24hTD).

Processeurs de traitement numérique du signal : à la rentrée 1997/98, j'ai introduit un nouveau cours en 5^e année du département EII portant sur les *Processeurs de traitement numérique du signal (DSP)* permettant ainsi aux étudiants de mettre en application et d'illustrer les notions théoriques acquises en cours de *Traitement du signal* des deux années précédentes.

Informatique scientifique : outre les encadrements de TP en 3^e et 4^e année du département EII, j'ai participé pendant longtemps (1997-2006), au cours de *Système d'Exploitation* et de la *programmation en langage C* en 3^e année du département MNT (12hTD).

Robotique-Productique : j'ai participé à la création et à l'encadrement des enseignements de Robotique-Productique en 5^e année GMA et EII.

Modules de sensibilisation : avec la mise en place du LMD à l'INSA de Rennes, je participe à l'élaboration et à l'encadrement de nouveaux modules d'enseignement en *Automatique* et en *Traitement du Signal* qui sont destinés aux étudiants de 2^e année de l'INSA.

Le tableau ci-dessous récapitule l'ensemble des matières principales dans lesquelles j'interviens sous forme de cours (CM), travaux dirigés (TD) ou travaux pratiques (TP).

	Niveau et forme d'enseignement										
Matière	2 ^e année		3 ^e année			4 ^e année			5 ^e année		
	TD	TP	CM	TD	TP	CM	TD	TP	CM	TD	TP
Signaux et systèmes			X	X							
Automatique	X	X		X	X						
Traitement du signal	X	X	X	X	X	X	X	X			
DSP									X		X
Langage C				X	X			X			
Moyenne annuelle de l'ordre									240 héqTD		

Tableau de synthèse des enseignement sur une année.

Outre ces enseignements traditionnels, je participe également à des enseignements comme :

- les *exposés technologiques*. Ceux-ci consistent à donner un article scientifique ou un sujet à un étudiant de 4^e année qui doit le résumer en anglais sur 6 pages et présenter son travail devant ses collègues.
- les *projets de langage C*. Je propose des sujets de réalisation et j'encadre des groupes de quatre étudiants dans la réalisation de leur projet de programmation en langage C. Cela consiste en la gestion et la réalisation d'un projet tout en coordonnant les relations humaines.
- les *stages de 4^e année*. J'encadre des stages des étudiants de la 4^e année du département EII.
- les *projets de fin d'études*. J'encadre, également, des projets de fin d'études des étudiants de la 5^e année pendant leur stage en entreprise.

V - Responsabilités administratives et scientifiques

V.1 - Responsabilités scientifiques

- Depuis 1994, j'anime le thème "Analyse et interprétation d'images" au sein du Groupe Image et Télédétection de l'IETR qui a eu quatre thèses soutenues et une vingtaine de projets de PFE ou de Master.
- Responsabilité scientifique, j'ai participé activement au développement dans le projet DRI/SIROS pour une étude d'explosion-fusion d'images pour la visualisation d'images SAR (2000-2001).

Référence : Marché CELAR n° 99 42 165 00 470 35 15

Partenaires : ENST Paris, INSA Rennes, ENST Brest, IRESTE Nantes, AQL Rennes et IPSIS Rennes

Coût total : 120 000 F (part de l'INSA)

- Responsabilité scientifique et réalisation dans le projet ADEME/ALLIGATOR pour la modélisation de profils de scies et de meule et la simulation de l'affûtage de scies à bois en vue de la fabrication d'une affûteuse à commande numérique (1992-1996).

Références : Conventions ALLIGATOR CR368 et 501/G004

Partenaires : ADEME, SA ALLIGATOR, SA TRANSIFLEX

Coût total : 60 000 F (par convention)

- Participation au projet ANR CAIMAN (Codage Avancé d'IMAgés et Nouveaux services) qui regroupe THALES Communications, l'équipe Image de l'IETR, le laboratoire XLIM de Poitiers, l'ETIS de Cergy-Pontoise

V.2 - Collaborations internationales

- collaboration avec le Laboratoire "Signaux et Systèmes" de la Faculté des Sciences de l'Ingénieur de l'Université de Sidi Bel-Abbes, Algérie depuis 2003. Cette collaboration de recherche nous a permis de publier plusieurs articles sur le

recalage d'images satellitaires. Cette collaboration devra conduire à une thèse en cotutelle.

- collaboration avec l'Université de Saint-Louis (Sénégal) pour l'encadrement d'une thèse de l'un de leur enseignant (thèse soutenue en 2007).
- conférence donnée à l'Université de Lomé (Togo), sur l'invitation du doyen de la fac. des sciences, sur le thème "Le Traitement d'Images et ses applications", 1999.

V.3 - Responsabilités diverses

V.3.1 - Révision d'articles et animation de sessions

- Je participe à l'évaluation des articles soumis pour publication dans des revues ou dans les actes de congrès comme CARI, ICIP, IASTED, EUSIPCO, IGARSS, revue Télédétection dont des centrales de révision d'articles :
 - Canadian Journal of Remote Sensing (<http://mc.manuscriptcentral.com/cjrs-jct>),
 - IEEE Transactions on Geoscience and Remote Sensing (<http://mc.manuscriptcentral.com/tgrs>)
 - IET Image Processing de The Institution of Engineering and Technology (<http://mc.manuscriptcentral.com/iet-ipr>)
- J'ai été coorganisateur et co-animateur du Séminaire "Quels DSP pour les applications de télécommunications et images ?", organisé par INSA de Rennes en collaboration avec Texas Instruments et ATEME le 07 octobre 2003
- J'ai été co-animateur de la session "Neural network II" à la conférence AIA'04 de IASTED à Innsbruck en février 2004.

V.3.2 - Responsabilités administratives et d'intérêt commun

- Membre du vivier du comité de sélection des sections 61 de l'INSA de Rennes, de l'IUT de Rennes et de l'université de Rennes 1.
- Chargé de missions aux Relations Internationales de l'INSA pour la Zone Afrique (depuis 2007)
- Membre élu du Conseil Scientifique de l'INSA de Rennes (depuis 2006).
- Membre élu du Conseil de Département EII de l'INSA de Rennes (2000-2002) et depuis 2005.
- Membre élu de la Commission des Spécialistes de la 61^e section de l'INSA de Rennes (depuis 1998). Je suis 2nd vice-président (depuis 2000)
- Membre élu de la Commission des Spécialistes de la 61^e section de l'Université de Rennes I (2000-2004)
- Membre élu du Comité d'Action Sociale de l'INSA (CASI) de Rennes (1999-2001)
- Participation à diverses commissions d'admission à Lyon ou à Rennes (depuis 1998) et à divers salons et/ou portes ouvertes à l'INSA
- Membre du groupe NTE de l'INSA de Rennes (depuis 2001)
- Participation à la gestion du matériel informatique PC du laboratoire (1990-1994).

VI - Travaux et publications

Cette section présente l'ensemble des publications issues de mes activités de recherche ou d'enseignement. On y trouvera des articles publiés dans des revues scientifiques ou dans les actes de conférences, des rapports de conventions ou autres travaux comme les manuels d'enseignement ou des documents pédagogiques. Ce bilan se décompose en deux parties : une partie dédiée aux travaux réalisés pendant la thèse et l'autre réservée aux réalisations post-doctorales.

VI.1 - Thèse et travaux associés

VI.1.1 - Thèse

[Th1] *Analyse fractale de textures naturelles dans un contexte multirésolution : application à la segmentation d'images multirésolution.*

Résumé :

Analyse de textures naturelles pour leur caractérisation par la dimension fractale basée sur le modèle du mouvement brownien fractionnaire. Après avoir montré, expérimentalement, que les textures naturelles étaient anisotropes vis-à-vis du paramètre fractal, l'attribut de textures est défini suivant différentes orientations.

En appliquant ces attributs à la segmentation de textures dans un contexte multirésolution basée sur les niveaux de la pyramide gaussienne, nous avons pu montrer que la combinaison de plusieurs résolutions donnait de meilleurs résultats de segmentation qu'en utilisant une seule résolution.

Date de soutenance : 12 février 1992 à l'INSA de Rennes.

Directeur de thèse : Joseph RONSIN

Jury :	Joseph RONSIN	Président
	René COLLOREC	Rapporteur
	Claude LABIT	Rapporteur
	Jean-Marc BOUCHER	Examineur
	Alain BRUNO	Examineur
	Véronique HAESE-COAT	Examinatrice
	Stéphane RABOISSON	Examineur

Mention : Très honorable

VI.1.2 - Conférences internationales avec comité de lecture [1 article]

- [CICL1] Kidiyo Kpalma, Alain Bruno and Véronique Haese-Coat, *Natural Texture Analysis in Multiscale context Using Fractal Dimension*, SPIE-Visual Communication and Image Processing, nov. 1991 à Boston, USA

VI.1.3 - Conférences nationales avec comité de lecture [2 articles]

- [CNCL1] Kidiyo Kpalma, Véronique Haese-Coat et Alain Bruno, *Analyse Pyramidale de Textures Naturelles à Travers la Dimension Fractale*, 13^e colloque de GRETSI, sept. 1991 à Juan-les-Pins, France
- [CNCL2] Kidiyo Kpalma, Alain Bruno et Joseph Ronsin, *Analyse de textures naturelles : application à la segmentation d'images*, Séminaire ORASIS, CNRS-PRC Homme-Machine, sept. 1991 à Saint-Malo, France

VI.1.4 - Conférences sans actes [1 Conférence]

- [COSA1] Kidiyo Kpalma, *Analyse fractale de textures dans un contexte multirésolution*, Séminaire Signal et Image, IRISA, nov. 1991.

VI.2 - Publication et travaux réalisés après la thèse

VI.2.1 - Revues internationales avec comité de lecture [7 articles]

- [RICL1] Ahsan A. Ursani, Kidiyo Kpalma and Joseph Ronsin, *Texture features based on local Fourier histogram: self-compensation against rotation*, Journal of Electronic Imaging 17, 03 (2008)
- [RICL2] Kidiyo Kpalma and Joseph Ronsin, *Turning Angle Based Representation for Planar Objects*, Electronics Letters / IEE Electronics Letters, Volume 43, No 10, pp. 561-563, (2007)
- [RICL3] Mingqiang Yang, Kidiyo Kpalma and Joseph Ronsin, *Affine invariance contour descriptor based on iso-area normalization*, Electronics Letters / IEE Electronics Letters, Volume 43, No 7, pp. 379-380, (2007)
- [RICL4] Mingqiang Yang, Kidiyo Kpalma and Joseph Ronsin, *Affine Invariance Contour Descriptor Based on the Equal Area Normalization*, IAENG International Journal of Applied Mathematics 36, 2 (2007)
- [RICL5] Kidiyo Kpalma and Joseph Ronsin, *Multiscale contour description for pattern recognition*, Pattern Recognition Letters, Vol.27, No.13, pp 1545-1559, 1 October 2006
- [RICL6] Youcef Bentoutou, Nasreddine Taleb, Kidiyo Kpalma and Joseph Ronsin, *An Automatic Image Registration for Applications in Remote Sensing*, IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, No. 9, pp 2127-2137, September 2005

- [RICL7] Olivier Thépaut, Kidiyo Kpalma and Joseph Ronsin, *Automatic registration of ERS and SPOT multisensor images in a data fusion context*, Elsevier's Forest Ecology and Management, Vol. 128, pp 93-100, 2000

VI.2.2 - Revues nationales avec comité de lecture [1 article]

- [RNCL1] Véronique Haese-Coat, Kidiyo Kpalma et Joseph Ronsin, *Apport comparatif de la texture en segmentation d'images SPOT*, Revue Photo-Interprétation, nov. 1994, Editions ESKA

VI.2.3 - Livre [1 livre]

- [LIVR1] Kidiyo Kpalma et Véronique Haese-Coat, *Traitement numérique du signal : Théorie et applications*, Collection Technosup, Editions Ellipses, Septembre 2003, ISBN 2-7298-1728-X, (2003)

VI.2.4 - Chapitres de livres [4 livres]

- [ChLI1] Mingqiang Yang, Kidiyo Kpalma and Joseph Ronsin, *A Survey of Shape Feature Extraction Techniques*, in Pattern Recognition, ISBN 978-3-902613-24-4, I-Tech Education and Publishing (Ed.), (2008)
- [ChLI2] Mingqiang Yang, Kidiyo Kpalma and Joseph Ronsin, *Contour Descriptor Based on Affine Invariance Normalization*, in Advances in Communication Systems and Electrical Engineering, ISBN 978-0-387-74937-2, Springer (Ed.) (2008)
- [ChLI3] Youcef Bentoutou, Nasreddine and Kidiyo Kpalma, *Evaluation and Assessment of Rigid, Affine, and Nonrigid Registration Techniques: Application to X-ray Angiography Images*, in Computer Vision Research Progress, ISBN 1-60021-992-6, Nova Science Publishers (Ed.), (2008)
- [ChLI4] Kidiyo Kpalma, Ronsin J., *An overview of advances of pattern recognition systems in computer vision*, in "Vision systems : segmentation and pattern recognition", Chapter 10; Obinata and Dutta (Eds.), I-Tech Education and Publishing, Vienna, 2007

VI.2.5 - Conférences internationales avec comité de lecture [26 articles]

- [CICL1] Miloud Chikr El-Mezouar, Nasreddine Taleb, Kidiyo Kpalma and Joseph Ronsin, *A New Intensity-Hue-Saturation Fusion Technique with Color Distortion Reduction for IKONOS Imagery*, IEEE 5th International Conference of Sciences of Electronic, Technologies of Information and Telecommunications (SETIT 2009), March 2009, Tunisia
- [CICL2] Yang Mingqiang, Kpalma Kidiyo, and Ronsin Joseph, *Shape Matching and Object Recognition Using Chord Contexts*, 12th International Conference Information Visualisation (IV08), 8-11 July 2008
- [CICL3] Kidiyo Kpalma, Mingqiang Yang and Joseph Ronsin, *Planar shapes descriptors based on the turning angle scalogram*, Lecture Notes in Computer

- Science (LNCS) - International Conference on Image Analysis and Recognition (ICIAR), 25-27 Juin 2008, Povoá de Varzim, Portugal (2008)
- [CICL4] Ahsan A. Ursani, Kidiyo Kpalma and Joseph Ronsin, *Improved Texture Descriptor with Features based on Fourier Transform*, International Multi Topic Conference (IMTIC '08), 11-12 April 2008, Mehran, Pakistan (2008)
- [CICL5] Ahsan A. Ursani, Kidiyo Kpalma and Joseph Ronsin, *Over-Segmentation of Feature Space for Initialisation of K-means Clustering*, World Congress on Engineering and Computer Science 2007 - International Conference on Machine Learning and Data Analysis, 24-26 October 2007, San Francisco, USA (2007)
- [CICL6] Ahsan A. Ursani, Kidiyo Kpalma and Joseph Ronsin, *Texture features based on Fourier transform and Gabor filters: an empirical comparison*, International Conference on Machine Vision "Machine Vision for Development", (ICMV'07), December 28-29, 2007, Islamabad, Pakistan (2007)
- [CICL7] Kidiyo Kpalma and Joseph Ronsin, *A Novel Multi-Scale Representation for 2-D Shapes*, Lecture Notes in Computer Science (LNCS) - International Conference on Image Analysis and Recognition (ICIAR), 22-24 August 2007, Montreal, Canada (2007)
- [CICL8] Youcef Bentoutou, Nasreddine Taleb, A. Bounoua, Kidiyo Kpalma, Joseph Ronsin, *Feature Based Registration of Satellite Images*, IEEE Xplore - 15th International Conference on Digital Signal Processing (DSP 2007), Royaume-Uni (2007)
- [CICL9] Mingqiang Yang, Kidiyo Kpalma and Joseph Ronsin, *An Affine Invariance Contour Descriptor Based on Filtered Enclosed Area*, The International MultiConference of Engineers and Computer Scientists 2007 (IMECS'07), March 21-23, Hong Kong, China (2007)
- [CICL10] Mingqiang Yang, Kidiyo Kpalma and Joseph Ronsin, *Scale-controlled area difference shape descriptor*, Document Recognition and Retrieval XIV Part of the IS&T/SPIE International Symposium on Electronic Imaging, San Jose, California, USA (2007)
- [CICL11] Mingqiang Yang, Kidiyo Kpalma and Joseph Ronsin, *A Shape Descriptor based on CSS*, Document Recognition and Retrieval XIV Part of the IS&T/SPIE International Symposium on Electronic Imaging, 28 January - 1 February 2007, San Jose Marriott and San Jose Convention Center San Jose, California, USA, (2007)
- [CICL12] Oumar Diop, Kidiyo Kpalma et Joseph Ronsin, *Approche texturale pour l'identification des nuages de poussières à partir d'images satellitaires infrarouges*, The Maghrebien Conference on Software Engineering and Artificial Intelligence (MCSEAI), (2006)
- [CICL13] Oumar Diop, Kidiyo Kpalma and Joseph Ronsin, *Textural approach for the detection of the dust clouds from infrared satellite*, Part of the SPIE Symposium on Asia-Pacific Remote Sensing, 13-17 November 2006, Goa, India. (2006)
- [CICL14] Youcef Bentoutou, Nasreddine Taleb, Kidiyo Kpalma and Joseph Ronsin, *A Feature-Based Approach to Automated Registration of Remotely Sensed*

Images, 2nd IEEE International Conference on information & Communication Technologies: from theory to applications (ICTTA'06), Damascus, Syria, April 24-28 2006

- [CICL15] Kidiyo Kpalma and Joseph Ronsin, *Contours Smoothing for Non-Occluded Planar Shapes Description*, International Conference on Image Processing (ICIP), 4 pages, Genova, Italy, 11-14 Sept. 2005
- [CICL16] Kidiyo Kpalma and Joseph Ronsin, *Planar Curve Smoothing for Pattern Recognition*, International Conference on Multispectral Image Processing and Pattern Recognition (MIPPR'03), Beijing, China, 20-22 October 2003
- [CICL17] Youcef Bentoutou, Nasreddine Taleb, Kidiyo Kpalma and Joseph Ronsin, *An invariant approach for automated image registration of multi-temporelles remotely sensed imagery*, The conference on Remote Sensing for Environmental Monitoring, GIS Applications, and Geology III, part of the Int. Symposium on Remote Sensing, Barcelona, Spain, 8-12 September 2003
- [CICL18] Kidiyo Kpalma and Joseph Ronsin, *A Multi-Scale curve smoothing for Generalised Pattern Recognition (MSGPR)*, ISSPA 2003, Paris, France, 1-4 July 2003
- [CICL19] Christophe Couturier, Kidiyo Kpalma and Joseph Ronsin, *DSP Teleoperation for Digital Signal Processing Teaching and Learning via Internet*, 14th International conference on DSP (DSP2002), Santorini, Greece, 1-3 July 2002
- [CICL20] Olivier Thépaut, Kidiyo Kpalma and Joseph Ronsin, *ERS SAR and SPOT Images Geocorrection and Registration in a context of Data Fusion*, Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, pp. 1040-1042, July 1998, Seattle, Washington, USA.
- [CICL21] Olivier Thépaut, Kidiyo Kpalma and Joseph Ronsin, *Multichannel remote sensing images registration in a consensual segmentation scheme*, The European Symposium on Remote Sensing, Conference on Image and Signal Processing for Remote Sensing IV, Vol. 3500, 21-27 Sept. 1998, Barcelona, Spain
- [CICL22] Olivier Thépaut, Kidiyo Kpalma and Joseph Ronsin, *Full Automatic Method of ERS and SPOT Images Registration in a Context of Data Fusion*, Proceedings of the First International Conference on Geospatial Information in Agriculture and Forestry, vol. 1, pp. 654-661, June 1998, Lake Buena Vista, Florida, USA
- [CICL23] Mustapha Assouil, Olivier Déforges and Kidiyo Kpalma, *Parallel implementation of a multiresolution image decomposition algorithm using A³ methodology*, SPIE'97 Annual Meeting, Conference on Parallel and Distributed Methods for Image Processing, 26-27 July 1997, San Diego, USA.
- [CICL24] Wei Li, Véronique Haese-Coat, Kidiyo Kpalma and Joseph Ronsin, *Texture Characterization by New Morphological Features: Application to SPOT Image Classification*, SPIE-2955 EUROPTO Conference on Image and Signal Processing for Remote Sensing III, 23-27 Sept. 1996, Taormina, Italie

- [CICL25] Benoît Ogor, Véronique Haese-Coat and Kidiyo Kpalma, *The cooperation of the mathematical morphology and region growing for remote sensing image segmentation*, Second European Satellite and Remote Sensing Symposium, Sept. 1995, Paris, France
- [CICL26] Kidiyo Kpalma, Véronique Haese-Coat and Joseph Ronsin, *A Hierarchical Multiresolution Texture Image Segmentation*, SPIE-Visual Communication and Image Processing, Nov. 1993, Boston, USA

VI.2.6 - Conférences nationales avec comité de lecture [6 articles]

- [CNCL1] Kidiyo Kpalma et Joseph Ronsin, *Description des Formes Planaires par Lissage Progressif de Contours Fermés*, 8^e Colloque Africain sur la Recherche en Informatique (CARI), 6-8 novembre 2006, Cotonou, Bénin.
- [CNCL2] Olivier Thépaut, Kidiyo Kpalma et Joseph Ronsin, *Recalage d'images de télédétection multi-source appliqué au filtrage multiscalaire et à la fusion de segmentation*, Colloque International La Télédétection Optique et Radar et la Géomatique pour la Gestion des Problèmes Environnementaux, 67^e Congrès de l'Association Canadienne-Française pour l'Avancement des Sciences (ACFAS), 18 pages, Ottawa, Ontario, Canada, mai 1999
- [CNCL3] Olivier Thépaut, Kidiyo Kpalma et Joseph Ronsin, *Géocorrection et Recalage d'images ERS et SPOT dans un contexte de fusion de données*, 4^e Colloque Africain sur la Recherche en Informatique (CARI), p. 891, octobre 1998 à Dakar, Sénégal.
- [CNCL4] Benoît Ogor et Kidiyo Kpalma, *Le Traitement Morphologique des Images Satellites : Une approche hiérarchique par région*, 3^e Colloque Africain sur la Recherche en Informatique (CARI), sept. 1996 à Libreville, Gabon.
- [CNCL5] Kidiyo Kpalma, *Caractérisation de textures par l'anisotropie de la dimension fractale*, 2^e Colloque Africain sur la Recherche en Informatique (CARI), octobre 1994 à Ouagadougou, Burkina Faso.
- [CNCL6] Kidiyo Kpalma et Véronique Haese-Coat, *Une Nouvelle Méthode de Segmentation d'Images de Textures Multirésolution*, Colloque de GRETSI, sept. 1993 à Juan-les-Pins, France

VI.2.7 - Conférences invitées [1 conférence]

- [CINV1] Kidiyo Kpalma, *Le Traitement d'Images et ses applications*, Conférence présentée le 18 février 1999 à l'Université de Lomé (Togo) sur l'invitation du Doyen de la Faculté des Sciences de cette Université.

VI.2.8 - Conférences sans acte [1 conférence]

- [CSA1] Kidiyo Kpalma, *TELEDITS : enseignement de DSP à distance*, Séminaire *Quels DSP pour les applications de télécommunications et images ?*, organisé par INSA de Rennes, Texas Instruments et ATEME, 07 octobre 2003

VI.2.9 - Rapports de conventions [7 rapports]

VI.2.9.1 - Explosion/Fusion pour la visualisation

- [RACO1] Kidiyo Kpalma, Joseph Ronsin et Olivier Adam, *Rapport Final (RP2)*, "Algorithmies pour l'aide à la détection, la reconnaissance et l'identification (DRI) des cibles militaires dans les données de l'observation. Poste 2 - Explosion / Fusion algorithmique - Marché CELAR n° 99 42 165 00 470 35 15", février 2001
- [RACO2] Olivier Thépaut, Kidiyo Kpalma et Joseph Ronsin, *Spécification de la maquette de démonstration (SMP2)*, Rapport de convention "Algorithmies pour l'aide à la détection, la reconnaissance et l'identification (DRI) des cibles militaires dans les données de l'observation. Poste 2 - Explosion / Fusion algorithmique - Marché CELAR n° 99 42 165 00 470 35 15", novembre 2000
- [RACO3] Olivier Thépaut, Kidiyo Kpalma et Joseph Ronsin, *Rapport Technique du Poste 2 (RTP2)*, Rapport de convention "Algorithmies pour l'aide à la détection, la reconnaissance et l'identification (DRI) des cibles militaires dans les données de l'observation. Poste 2 - Explosion / Fusion algorithmique - Marché CELAR n° 99 42 165 00 470 35 15", septembre 2000

VI.2.9.2 - Modélisation de dents de scies

- [RACO4] Kidiyo Kpalma et Joseph Ronsin, *Génération de la trajectoire de meule pour le contrôle d'une machine à commandes numériques dans un système de d'affûtage de scies à bois*, Rapport de convention, 3^e convention INSA/S.A. ALLIGATOR (étude et développement), N° 501/G004, mars 1996
- [RACO5] Kidiyo Kpalma et Joseph Ronsin, *Modélisation et numérisation de profils de dentures de scies à bois*, Rapport de convention, 2^e convention INSA/ADEME/S.A. ALLIGATOR (étude et développement), N° CR368 : deuxième volet, mai 1993
- [RACO6] Kidiyo Kpalma et Joseph Ronsin, *Modélisation et numérisation de profils de dentures de scies à bois*, Rapport de convention, 1^{ère} convention INSA/ADEME/S.A. ALLIGATOR (étude de faisabilité et réalisation), N° CR368 : premier volet, juin 1992

VI.2.9.3 - Assistance à la Gestion de l'Environnement : projet AGER

- [RACO7] Benoît Ogor, Véronique Haese-Coat, Kidiyo Kpalma et Joseph Ronsin, *Segmentation hiérarchique d'images satellites par approche région et morphologie mathématique*, Rapport, Projet CBI-AGER : CR2A (Rennes), ENSTB (Brest), ENSSAT (Lannion), ORSTOM (Lannion) et COSTEL (Rennes), Réf. Région Bretagne, Opération n° 1213893, Programme n° 691, 1995, sept. 1995

VI.2.10 - Logiciels [2 logiciels]

- [LOGI1] Christophe Couturier, Kidiyo Kpalma et Joseph Ronsin, Logiciel DspEdit, pour la téléopération de DSP dans le cadre du projet TELEDITS, juin 2001

- [LOGI2] Kidiyo Kpalma, Logiciel ALIGATOR, pour la génération de trajectoire de meule par simulation de l'affûtage. Logiciel commercialisé par la S.A. ALLIGATOR avec les affûteuses à commandes numériques, Dernière version en septembre 1997

VI.2.12 - Documents pédagogiques

[POLY1] ***Théorie et traitements des signaux***

Coauteur : .. V. COAT

Public concerné :3^e année EII

Nombre de pages :154

Contenu
<ul style="list-style-type: none"> ▪ Signaux : définition, classification, représentation, ▪ Signaux déterministes, ▪ Signaux aléatoires, ▪ Filtrage linéaire, ▪ Modulation, démodulation, ▪ Détection et estimation.

[POLY2] ***Signaux et systèmes***

Coauteur : ..

Public concerné :3^e année EII

Nombre de pages :215

Contenu
<ul style="list-style-type: none"> ▪ Généralités sur les signaux, ▪ Généralités sur les systèmes, ▪ Réponse d'un système linéaire à une excitation quelconque, ▪ Transformations de Fourier , de Laplace, ▪ Fonction de transfert d'un système, ▪ Etude de stabilité d'un système et cartographie de ses pôles, stabilité des systèmes munis d'une boucle de réaction.

[POLY3] ***Traitement Numérique du Signal***

Coauteur : .. V. HAESE-COAT

Public concerné :4^e année EII

Nombre de pages :172

Contenu
<ul style="list-style-type: none"> ▪ Echantillonnage et Quantification, ▪ Transformée de Fourier Discrète, ▪ Transformée en Z, ▪ Les Transformations Unitaires, ▪ Communications Numériques et Modulations Numériques, ▪ Filtrage Numérique

[POLY4] ***Processeurs de Traitement Numérique du Signal***

Coauteur : ..

Public concerné :5^e année EII

Nombre de pages :216

Contenu
<ul style="list-style-type: none"> ▪ Vue d'ensemble de l'architecture des DSP de la famille TMS320, ▪ Etude de l'architecture fonctionnelle des DSP TMS320C5X, ▪ Modes d'adressage : adressage direct, adressage indirect, adressage immédiat et adressage circulaire, adressage bit-reverse, ▪ Outils de développement : étude d'un DSK et assembleur des 'C5X.

[POLY5] Programmation en langage C

Coauteur : .. C. MENENDEZ

Public concerné :3^e année MNT

Nombre de pages :95

Contenu

- Structure d'un programme en langage C
- Types de données en langage C
- Entrées-Sorties standard
- Structures alternatives ou instructions conditionnelles
- Instructions de boucles
- Tableaux et pointeurs
- Fonctions en langage C
- Fichiers en langage C

[POLY6] Modules de Sensibilisation**Traitement du signal**

Coauteurs :M. BABEL, V.

HAESE-COAT et J.-N. PROVOST

Public concerné :2^e année

Nombre de pages :23

Contenu

- Présentation de la discipline
- TP découverte

AutomatiqueCoauteurs :E. BAZIN et
M. JULLIEREPublic concerné :2^e année

Nombre de pages :18

- Présentation de la discipline
- TP découverte

VI.3 - Bilan des travaux dans la période

Le tableau ci-dessous récapitule l'ensemble des publications faites pendant la période post-doctorale de 1993 à 2008.

On notera que la production des articles scientifiques est fortement modulée par le développement de projets de conventions dont les rapports sont comptabilisés dans la rubrique "Rapports" : les périodes de faible production d'articles correspondent aux périodes de fortes activités sur les contrats. Dans ce bilan, je n'ai pas considéré les documents pédagogiques comme les polycopiés de cours, mais je dois signaler que leur production a, comme les conventions, influé sur la production des articles.

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Revue		1						1					1	1	3	1
Conf. actes	2	1	1	2	1	4	1			1	3		1	4	7	3
Chapitres de livres															1	3
Livres											1					
Rapports	1		1	1				2	1							
Total	3	2	2	3	1	4	1	3	1	1	4	0	2	5	11	6
Nombre de revues sur les 4 dernières années :									6 revues internationales							

Tableau récapitulatif de la production scientifique : la rubrique Revues regroupe les revues nationales et internationales. De même, la rubrique Conférences avec actes regroupe les conférences nationales et internationales.

Deuxième partie

Rapport de synthèse

VII - Analyse multiéchelle de contours de formes planaires

VII.1 - Introduction

Cette section est principalement dédiée à l'analyse multiéchelle de contours. Elle a pour objectif de présenter, plus en détails, l'une de nos méthodes d'analyse multiéchelle pour la description d'objets planaires.

Le volume sans cesse croissant des données électroniques (images, texte...) induit une difficulté à retrouver rapidement et facilement les informations recherchées. Il devient donc nécessaire, voire obligatoire, de trouver des solutions automatiques ou semi-automatiques pour la recherche de l'information dans de grandes bases de données. Dans cette situation les mots-clés comme ***Content-Based Image Retrieval (CBIR)***, ***Query by Image Content (QBIC)***, ***Content-Based Visual Information Retrieval (CBVIR)*** ou ***recherche d'images par le contenu*** deviennent de plus en plus courants et prennent tout leur sens dans notre quotidien. Toutes ces différentes expressions ont pour objectif de définir un système générique permettant d'explorer de grandes bases d'images et d'en récupérer celles qui sont similaires à la requête en utilisant des critères basés sur le contenu des images [1, 3-9]. Par opposition aux systèmes traditionnels qui utilisent la recherche contextuelle par des mots-clés (ex. Google) [2], un système du type CBIR s'appuie sur les caractéristiques "intrinsèques" contenues dans l'image requête pour réaliser la recherche. Les caractéristiques les plus couramment utilisées sont : la texture, la couleur, le pourcentage de couleur, les formes ou une combinaison de plusieurs de ces caractéristiques (ex. Image-Seeker[®] de LTU Technologies, PhotoBook de MIT Media Laboratory) [3-5].

En guise d'exemple pratique, on peut citer Flickr qui est un site de partage de photos et de vidéos. Il dit avoir hébergé plus de 3 milliards de photos en novembre 2008. Comment peut-on s'y retrouver? Grâce aux applications comme retrievR et Xcavator on peut effectuer des recherches d'images à partir d'une image requête ou tout simplement à partir d'un dessin à main levée [10].

Conscients de cette situation, nous avons entrepris le développement d'un nouvel axe de recherche, au sein du groupe Image et Télédétection de l'IETR, afin de proposer des solutions pour la description et la reconnaissance des formes. Nous chercherons donc à approfondir et améliorer nos méthodes d'analyse pour mettre en place des descripteurs efficaces et robustes pour l'indexation des images pour des applications de recherches par le contenu ou par similarité tout en se plaçant dans le cadre des standards existants comme MPEG7. A moyen terme, l'intégration et l'adaptation de tels descripteurs dans un processus plus complet permettront de mettre en place un système de suivi d'objets dans une séquence d'images.

Dans le contexte de la reconnaissance automatique des formes, il est nécessaire de disposer d'un algorithme simple, rapide et efficace surtout lorsqu'on a besoin d'explorer une grande base de données.

Un schéma générique (voir Fig.VII.9) de la reconnaissance des formes comporte deux phases essentielles :

- la **phase d'analyse** où l'on extrait les attributs (l'indice) qui caractérisent chaque objet ou chaque modèle (classe) de la base dans la phase d'apprentissage (❶),
- la **phase de reconnaissance** où l'on va chercher à reconnaître un objet en le comparant au modèles de la base en se basant sur les attributs définis dans la phase d'analyse (❷).

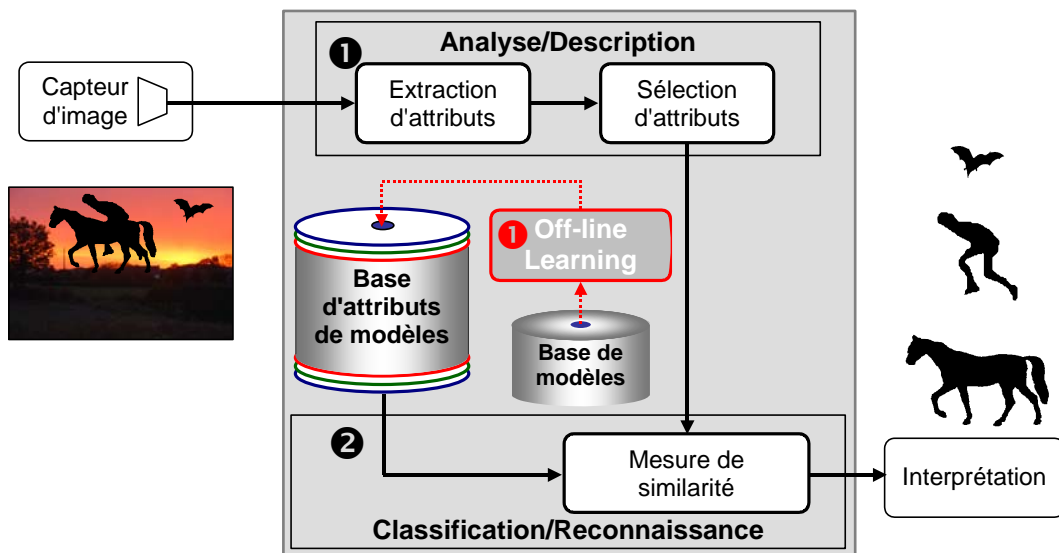


Fig.VII.9 : Un schéma générique d'un système de reconnaissance de formes

La littérature propose plusieurs types de méthodes sur le sujet, mais la définition et la sélection des attributs, problème crucial, reste toujours très complexe et exige un grand nombre d'attributs [Mokhtarian96, Sossa00] : surface de la forme, facteur d'élongation, moments invariants, textures, nombre de trous, ...

Dans cette section, nous proposons une nouvelle méthode d'analyse (la méthode *MSGPR* : *Multi-Scale curve smoothing for Generalised Pattern Recognition*) [Kpalma03, Kpalma06] permettant d'extraire des attributs caractéristiques intéressants. Cette méthode est basée sur un lissage multiéchelle par filtrage progressif du contour d'une forme plane. De par sa structure, cette méthode est simple de mise en œuvre et s'avère invariante aux translations et aux rotations et présente une résistance au changement d'échelle.

Pour commencer, nous présentons le principe de la méthode MSGPR ainsi que ses fondements. Après un aperçu de certaines propriétés de cette méthode, nous proposons quelques attributs intéressants que l'on peut extraire pour la caractérisation des formes planaires ; nous développons ensuite quelques exemples d'application en guise d'évaluation de la méthode. Pour finir, nous tirons les conclusions de cette étude et présentons les perspectives pour la suite des travaux sur ce thème.

VII.2 - Principe de la méthode MSGPR

VII.2.1 - Préliminaires

La méthode MSGPR est une méthode basée sur les contours fermés de la forme à analyser et suppose donc un traitement préalable de détection de contours qui ne rentre pas dans le cadre de cette étude. C'est une méthode de la famille des méthodes d'analyse multiéchelle ou d'espace échelle comme la méthode CSS (Curvature Scale Space ou espace de courbure multiéchelle) [Mokhtarian92, Mokhtarian96, Matusiak99, Lindeberg94]. Cette dernière applique un filtrage gaussien d'écart-type σ au contour paramétré par le paramètre curviligne u . En réitérant cette opération pour des valeurs croissantes de σ , elle extrait les points d'inflexion de la courbure du contour filtré. En localisant tous les points d'inflexion de la courbure euclidienne à chaque paramètre σ du filtrage, dans l'espace (u, σ) , on obtient une série de courbes représentatives du contour analysé. La figure VII.10 illustre un contour (en VII.10.a) et les courbes de ses points d'inflexion (en VII.10.b).

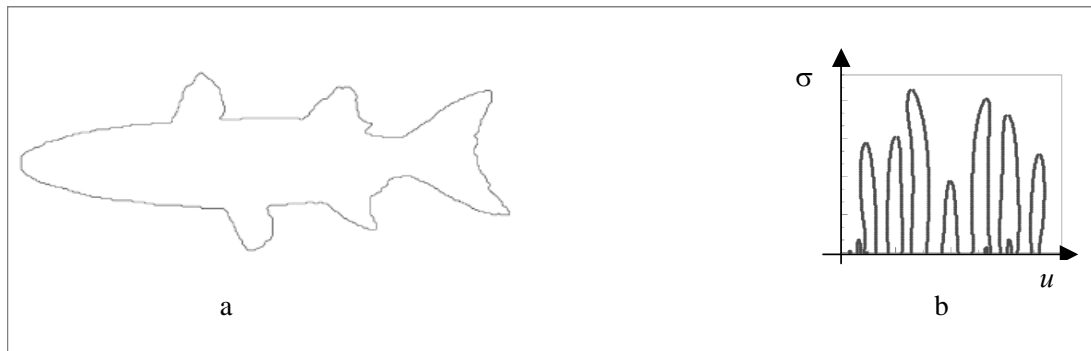


Fig.VII.10 : Un contour et sa carte CSS

Cette méthode met, donc, en jeu les opérations de filtrage, de calcul de courbure et de détection de points d'inflexion.

Comme on le verra plus loin, dans le cas de notre méthode, les seules opérations "lourdes" sont le filtrage et la détection de points d'intersection.

VII.2.2 - Structure de la méthode MSGPR

Cette méthode peut se structurer, simplement en quatre tâches principales. Comme le montre la figure VII.11 les différentes étapes dans sa mise en œuvre sont :

- 1) décomposition du contour original C_0 en deux fonctions $x(u)$ et $y(u)$ qui sont fonction du paramètre curviligne u ,
- 2) lissage progressif des deux fonctions par filtrage passe-bas gaussien d'écart-type σ ,

- 3) ajustement d'échelle et la reconstruction du contour lissé $C_{GC}(\sigma)$,
- 4) extraction des points d'intersection entre le contour original C_0 et le contour lissé $C_{GC}(\sigma)$ puis la génération de la carte CPI.

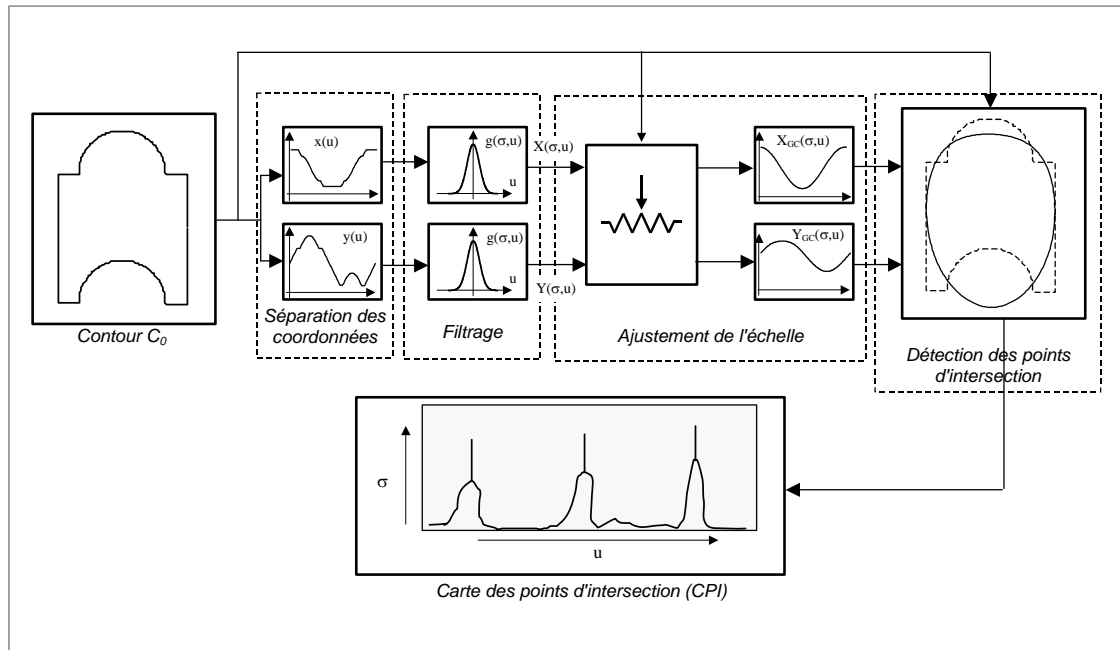


Fig.VII.11 : Structure de la méthode MSGPR

VII.2.3 - Prétraitement

Avant d'appliquer cette méthode à un quelconque contour celui-ci doit subir un prétraitement de normalisation qui consiste en trois opérations :

- 1) ré-échantillonnage du contour afin d'obtenir 360 points équidistants. Le choix de 360 points est arbitraire et se justifie, simplement, par le fait que dans le cas limite d'un cercle, on obtient un écart angulaire de 1° entre deux points consécutifs. De plus, comme on le verra plus loin, ce choix permet une lecture directe de l'angle de rotation de la forme sur la carte CPI.
- 2) changement d'échelle pour que la distance inter-points soit de l'ordre de l'unité. Etant donné que nous travaillons sur des formes représentées par une image, ce traitement permet d'obtenir des valeurs entières correspondant aux pixels dans une image : le point (x,y) sera ainsi représenté par le pixel (x,y) . Cette opération correspond à une normalisation, car quelle que soit la forme cet ajustement d'échelle permet d'avoir une distance de l'ordre de l'unité entre deux points consécutifs.
- 3) rotation permettant d'aligner la forme selon son axe principal le long de l'axe des abscisses. Cela permet, ainsi, d'avoir les mêmes fonctions x et y quelle que soit l'orientation de l'objet et, par conséquent, les mêmes résultats de filtrage. En effet comme le montre la figure VII.12, les fonctions x et y diffèrent en fonction de l'orientation de la forme originale : ceci conduit à un contenu spectral différent se traduisant par des contours lissés avec des points d'intersection totalement différents.

Ce prétraitement nous donne ainsi le contour original d'entrée C_0 . Dans la suite, les expressions "contour C_0 " ou "contour original" désigneront le contour prétraité.

VII.2.4 - Séparation des coordonnées

Le contour original C_0 est décrit par un ensemble de points représentés par leurs coordonnées (x, y) centrées (autour du centre de gravité G). La première opération consiste à paramétrer le contour à l'aide de la coordonnée curviligne u . Chacun de ses points est alors représenté par ses coordonnées paramétriques $(x(u), y(u))$. Ensuite, le contour est séparé en deux fonctions $x(u)$ et $y(u)$ correspondant, respectivement, aux abscisses et aux ordonnées des points du contour. Le paramètre u est alors normalisé par rapport à la longueur du contour : ainsi le domaine de validité de u varie de 0 à 2π (ou de 0° à 360°). Le sens d'évolution de ce paramètre est le sens trigonométrique avec son origine placée à un point P_0 déterminé au préalable. Ce point doit être unique et stable quelle que soit l'orientation de la forme.

La figure VII.12 montre un exemple de fonctions $x(u)$ et $y(u)$ avant et après rotation de 45° . Sur cette figure, le point initial est le point du contour qui est le plus à droite sur l'axe des abscisses. Il apparaît clairement sur cette figure que les fonctions $x(u)$ et $y(u)$ sont totalement différentes par la rotation alors que le contour lui-même reste topologiquement identique. En effet la rotation (dans le sens trigonométrique) d'un contour de coordonnées $(x(u), y(u))$ d'un angle θ donne un contour de coordonnées $(x(\theta, u), y(\theta, u))$ à l'aide de la relation :

$$\begin{pmatrix} x(\theta, u) \\ y(\theta, u) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x(u) \\ y(u) \end{pmatrix} = \begin{pmatrix} x(u)\cos(\theta) - y(u)\sin(\theta) \\ x(u)\sin(\theta) + y(u)\cos(\theta) \end{pmatrix} \quad (5)$$

L'équation (5) montre bien que l'abscisse $x(\theta, u)$ ne dépend pas uniquement de $x(u)$ mais aussi de $y(u)$; de même, $y(\theta, u)$ ne dépend pas uniquement de $y(u)$ mais aussi de $x(u)$. Afin de contourner ce problème, il est important de choisir le point initial de façon qu'il soit unique et stable : nous proposons donc une solution pour la détermination du point initial.

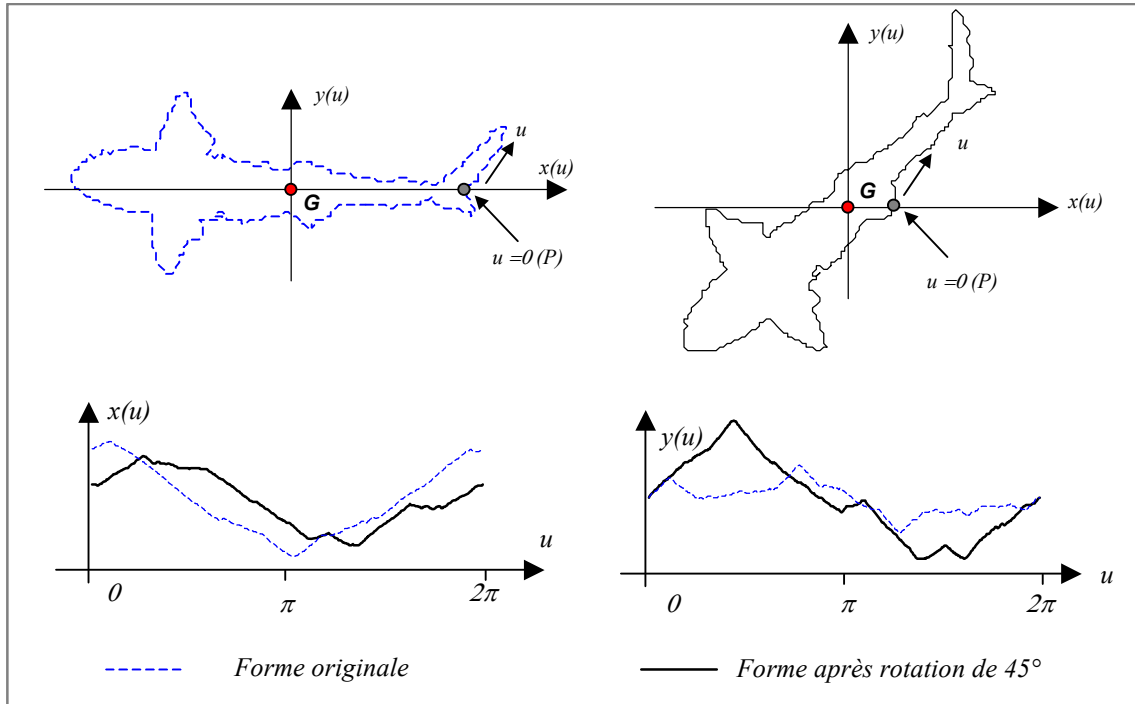


Fig.VII.12 : Un contour dans l'espace de travail

VII.2.5 - Détermination du point initial

Une méthode conduisant à un point initial suffisamment stable pour cette opération est celle basée sur les moments d'inertie. Après avoir déterminé les moments d'inertie, on définit les vecteurs d'inertie qui nous donne ainsi un référentiel "inertiel" (G, x_I, y_I) où l'axe Gx_I correspond à l'axe principal majeur et Gy_I à l'axe principal mineur qui sont orthogonaux dans un plan. Dans ce référentiel, nous choisissons le point P_0 comme étant le point $(x(u), y(u))$ de C_0 le plus éloigné du centre de gravité et appartenant à l'axe principal d'inertie Gx_I .

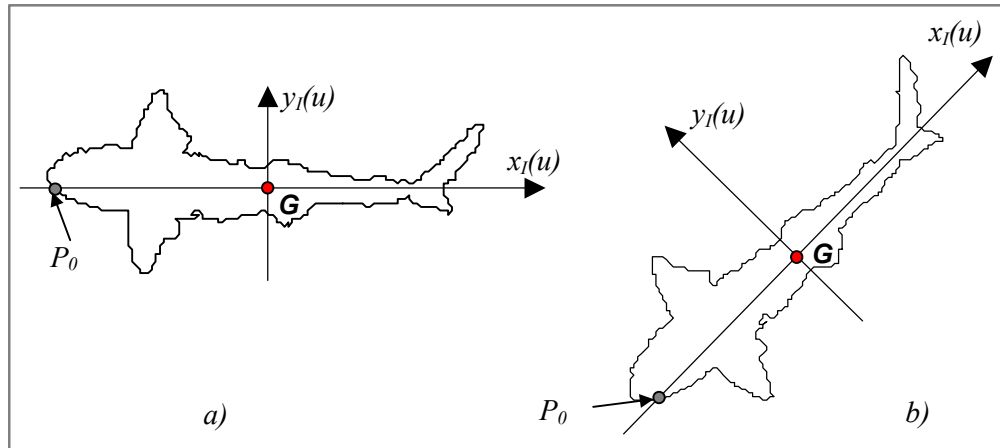


Fig.VII.12 : Exemples de contours dans le référentiel "inertiel" : b) forme tournée de 45° par rapport à celle de a)

Une fois cela fait, on peut choisir, soit, de rester dans le référentiel "inertiel" ou soit de passer dans le référentiel cartésien traditionnel. Pour des raisons de lisibilité, nous avons choisi de passer dans le référentiel traditionnel. Pour cela, on détermine l'angle trigonométrique θ entre l'axe GP_0 et l'horizontal puis on applique une rotation d'angle $-\theta$ pour ramener le point P_0 sur l'axe horizontal et le plus à droite du centre de gravité comme le montre la figure VII.13 ci-dessous.

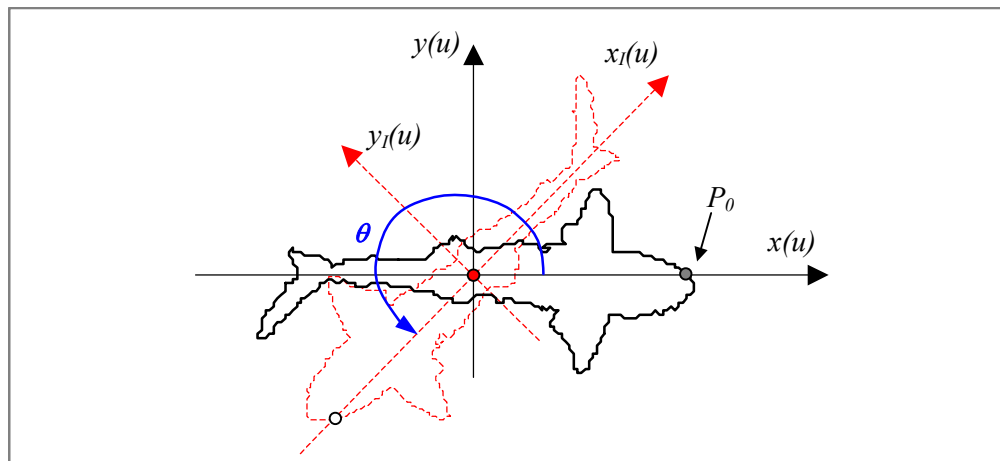


Fig.VII.13 : Exemple de repositionnement dans le référentiel (G, x, y)

Cette figure montre, également, le référentiel de travail dont l'origine est placée au centre de gravité G du contour. On notera que pour un contour régulier comme le cercle, du fait de l'échantillonnage régulier du paramètre u normalisé à 2π , ce paramètre peut être considéré comme un angle avec un écart angulaire de 1° entre deux points consécutifs.

VII.2.6 - Lissage de contours

Le lissage du contour est ensuite réalisé à l'aide d'un filtrage passe-bas par réduction progressive de la largeur de bande. Comme dans la méthode CSS le filtrage est basé sur le filtre gaussien de réponse impulsionnelle $g(\sigma, u)$ où σ est l'écart-type du filtre^{*}.

Pour un filtre gaussien de réponse impulsionnelle $g(\sigma, u)$ définie par

$$g(\sigma, u) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{u^2}{2\sigma^2}} \quad (6)$$

avec σ comme écart-type, la fonction de transfert $G(\sigma, f)$ est donnée par l'équation (7)

$$G(\sigma, f) = e^{-2\pi^2\sigma^2 f^2} \quad (7)$$

La fréquence de coupure f_c à $-3dB$ d'un tel filtre est donnée par $f_c = \beta / \sigma$ où $\beta = \sqrt{\ln(2)} / 2\pi = 0.1325$. La largeur de bande de ce filtre est inversement proportionnelle au paramètre σ . Il est donc évident que l'augmentation du paramètre σ conduit à la réduction de la bande passante. Dans ces conditions, le filtre coupe de plus en plus bas quand on augmente progressivement le paramètre σ . Ainsi, lorsque σ tend vers l'infini, la réponse $X(\sigma, u)$ (ou $Y(\sigma, u)$) du filtre tend vers la composante continue du signal considéré : G_x et G_y , pour x , et pour y respectivement où G_x et G_y sont les coordonnées du centre de gravité \mathbf{G} du contour.

La figure VII.14.a illustre le résultat d'un filtrage d'un contour pour $\sigma=30$ et $\sigma=180$. Sur cette figure sont représentés les signaux x et y originaux, les réponses correspondantes du filtre ainsi que les contours reconstruits correspondants. Il y apparaît clairement que les contours lissés C_σ ne sont pas à la même échelle que l'original. Il serait donc absurde de chercher à déterminer les points d'intersection de tels contours avec le contour original : certaines valeurs de sigma (ex. $\sigma=180$) ne donnent aucun point d'intersection. Pour pouvoir obtenir des points d'intersection cohérents, il faut procéder à un ajustement d'échelle afin de mettre les deux contours à la même échelle.

Un ajustement adéquat de l'échelle appliqué aux contours lissés C_σ nous donne des contours $C_{GC}(\sigma)$ qui font apparaître des points comme on peut le voir sur la figure VII.14.b.

^{*} Le choix du filtre gaussien est motivé par le fait que nous souhaitons, dans un premier temps, nous mettre dans la même situation que la méthode CSS afin de pouvoir mener une étude comparative dans les mêmes conditions. Il est donc clair qu'en dehors de cette raison, l'on peut explorer d'autres types de filtres passe-bas pour, éventuellement, sélectionner le plus performant : une autre étude a été amorcée pour appliquer un filtre de lissage basé sur les ondelettes. Mais il est important de noter également que le filtre gaussien offre des propriétés particulières qui ont été reprises dans [Lindeberg94, pp31-60]

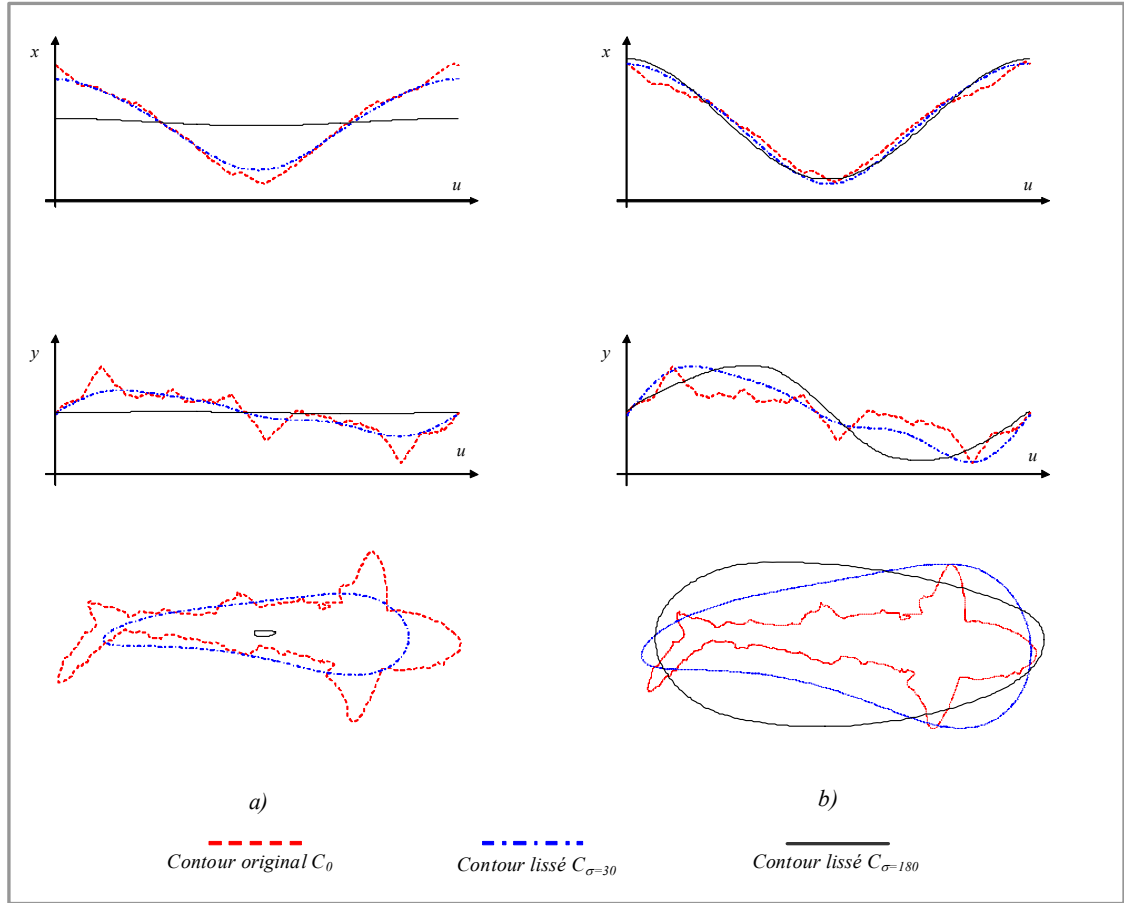


Fig. VII.14 : Signaux filtrés et contour lissé correspondant

VII.2.7 - Ajustement d'échelle

Comme il est précisé plus haut, le but de ce traitement est de transformer le contour original C_0 et le contour C_σ issu du filtrage de telle sorte qu'ils soient à la même échelle et qu'ils aient, éventuellement, des points d'intersection. En effet, comme il a déjà été mentionné, plus σ est grand et plus le contour C_σ se rétrécit (tendant vers le centre de gravité G). Pour que ce contour puisse avoir des points d'intersection avec le contour original, il est donc nécessaire de leur appliquer une opération de mise à l'échelle.

Si Z représente une séquence paramétrée par u et dont la moyenne est Z_G . Le coefficient A_Z étant le paramètre d'ajustement d'échelle permettant de ramener la séquence Z à l'échelle désirée : on obtient alors une nouvelle séquence Z_{GC} . Plusieurs solutions peuvent être envisagées en considérant l'équation suivante :

$$Z_{GC}(u) = A_Z[Z(u) - Z_G] + Z_G \quad (8)$$

Ici, nous ne présentons que la transformation appliquée à la séquence X des abscisses du contour lissé C_σ en sachant que le traitement est équivalent pour les ordonnées Y .

Tout d'abord, on calcule le coefficient AI_x des abscisses du contour original C_0 par l'expression suivante :

$$AI_x = \frac{1}{N} \sum_{n=1}^N |x(u_n) - x_G| \quad (9)$$

où les u_n sont les $N=360$ échantillons du paramètre u . En appliquant le même traitement aux abscisses du contour lissé, on obtient le coefficient AO_x :

$$AO_x = \frac{1}{N} \sum_{n=1}^N |X(\sigma, u_n) - X_G(\sigma)| \quad (10)$$

On définit, ainsi, le coefficient d'ajustement d'échelle A_x

$$A_x = \frac{AI_x}{AO_x} \quad (11)$$

On peut alors générer les abscisses du contour lissé et remis à l'échelle par

$$X_{GC}(\sigma, u_n) = A_x [X(\sigma, u_n) - X_G(\sigma)] + X_G(\sigma) \quad (12)$$

En procédant de la même façon, on détermine le coefficient A_y pour calculer la séquence $Y_{GC}(u, \sigma)$ des ordonnées à l'aide de l'équation (13) suivante et ainsi, on peut reconstruire le contour lissé et remis à l'échelle $C_{GC}(\sigma)$.

$$Y_{GC}(\sigma, u_n) = A_y [Y(\sigma, u_n) - Y_G(\sigma)] + Y_G(\sigma) \quad (13)$$

La figure VII.15.b montre un exemple (pour $\sigma=50$) des trois types de contours qui entrent en jeu dans cette méthode : C_0 , C_σ et $C_{GC}(\sigma)$.

Remarque

La définition donnée, ici, pour les coefficients AI_x et AO_x revient à normaliser les contours par rapport à l'écart moyen de leurs coordonnées. Il peut être envisagé d'utiliser d'autres coefficients comme l'écart maximal, l'écart quadratique moyen, l'amplitude crête à crête...

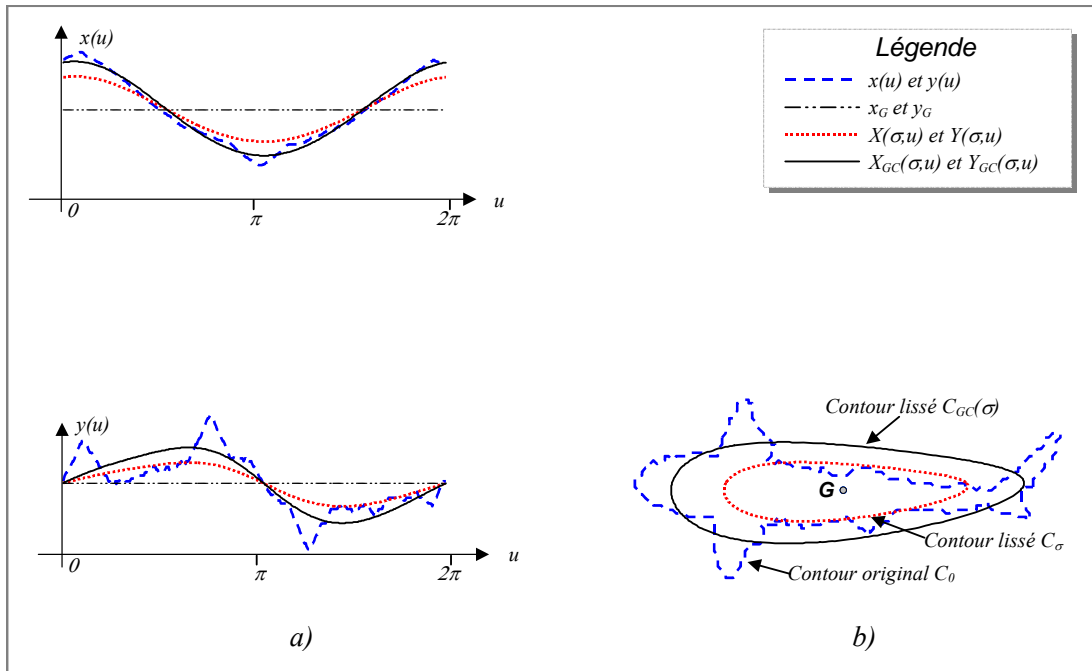


Fig.VII.15 : a) les fonctions x , y et leurs versions lissées et mises à l'échelle, b) contour d'entrée, contour lissée et le contour lissé mis à l'échelle

VII.2.8 - Définition de la carte CPI

Une fois le contour lissé mis à l'échelle du contour original, il ne reste plus qu'à extraire les points d'intersection entre le contour original C_0 et le contour lissé $C_{GC}(\sigma)$.

Pour ce faire, nous proposons une méthode permettant de définir la Carte des Points d'Intersection (CPI) ou IPM (Intersection Points Map) qui servira de base pour l'extraction d'attributs caractéristiques pour la description des formes. Cette carte est une fonction à deux variables σ et u et est définie de la façon suivante :

$$CPI(u, \sigma) = \begin{cases} 0 \\ \text{ou} \\ 1 \end{cases} \quad (14)$$

- $CPI(u, \sigma) = 0$ (noir sur l'image) si le point $(x(u), y(u))$ est un point d'intersection entre le contour original et le contour lissé et remis à l'échelle $C_{GC}(\sigma)$,
- $CPI(u, \sigma) = 1$ (blanc sur l'image) si le point $(x(u), y(u))$ n'est pas un point d'intersection des deux contours.

Il s'agit de prendre une valeur de σ correspondant à une largeur de bande B du filtre, d'appliquer le filtre aux deux signaux $x(u)$ et $y(u)$, de générer le contour $C_{GC}(\sigma)$ puis d'extraire les points d'intersection de ce contour avec le contour original et de marquer leur localisation dans la carte CPI à l'aide de la fonction CPI comme indiqué par l'équation (14) ci-dessus.

Pour $\sigma=0$, le filtre gaussien est un filtre passe-tout et alors tous les points du contour sont des points d'intersection et donc des points CPI.

En faisant croître la valeur du paramètre σ et en réitérant cette procédure, on obtient des filtres passe-bas qui coupent de plus en plus bas générant, ainsi, un lissage progressif de plus en plus fort. Les courbes ainsi générées tendent progressivement vers une courbe convexe (voir Fig.VII.16.a) et le nombre de points d'intersection tend vers une limite N_L . La trace laissée par un point d'intersection tend quant à elle vers une ligne plus ou moins verticale. Cette ligne peut devenir une bande plus ou moins large dans le cas où le contour lissé $C_{GC}(\sigma)$ est tangent au contour original C_0 .

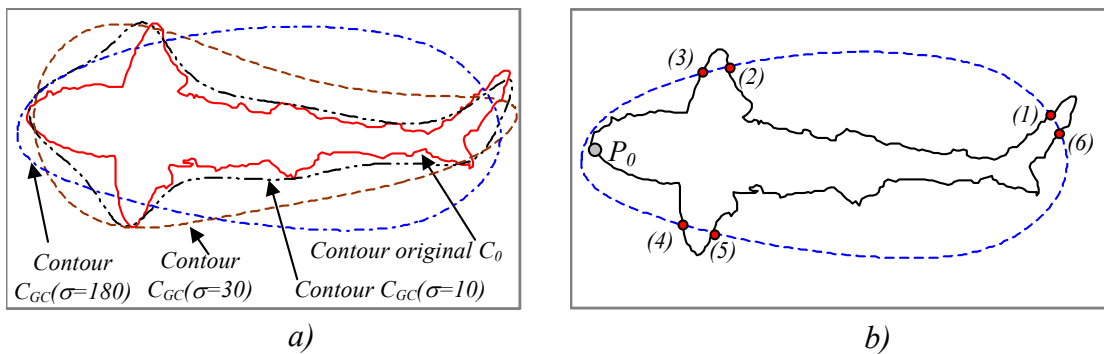


Fig.VII.16 : a) Contour original et deux versions lissées et b) : Points d'intersection entre le contour original et le contour lissé

Sur la figure 16.b on peut repérer facilement les points d'intersection dans le cas d'un filtre gaussien d'écart-type $\sigma=180$.

La figure VII.17 suivante montre un exemple de carte CPI obtenue à partir du contour original de la figure 16.b précédente. Sur cette carte, on distingue nettement la localisation des points d'intersection limites marqués (1) à (6) dans l'espace (u, σ) et qui correspondent aux points d'intersection repérés sur la figure 16.b par les mêmes chiffres.

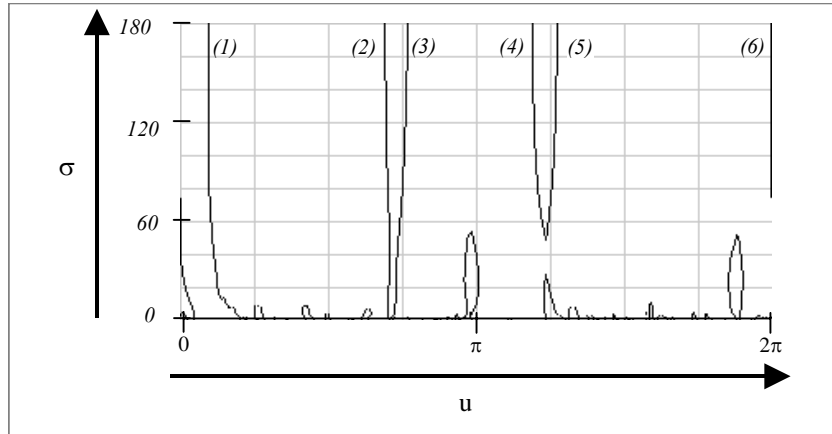


Fig.VII.17 : Exemple de carte CPI

On peut, également, noter l'apparition des structures comme celles notées (A) et (B) et (C) sur la figure VII.18. Ces structures caractérisent les portions convexes ou concaves d'un contour. Il s'agit des portions du contour qui, à partir d'un niveau de filtrage sont totalement disjointes du contour lissé.

Dans certains cas, un ou plusieurs points de cette portion réapparaissent comme points d'intersection : c'est le cas de la portion (C) de la figure VII.18 qui réapparaît sous l'ensemble des points (4) et (5) sur la figure VII.18 après avoir disparu pour une plage de σ .

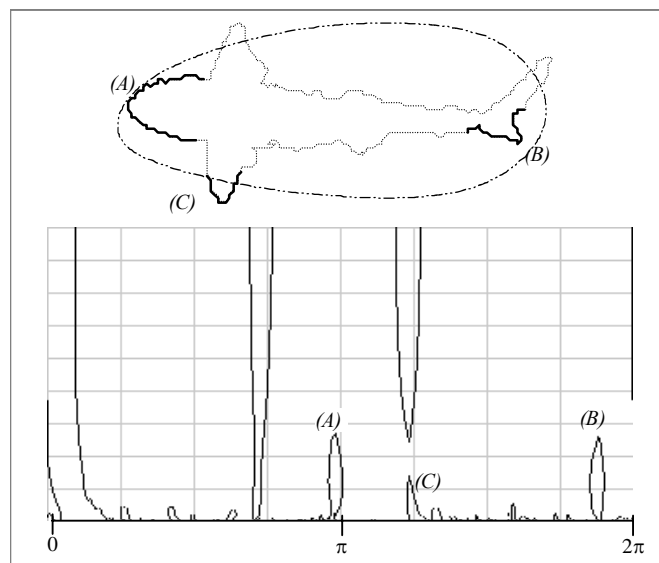


Fig.VII.18 : Caractérisation de portions convexes dans la carte CPI

En conclusion, on peut dire que la carte CPI nous révèle les deux informations suivantes sur le contour d'entrée :

- les points limites (en général, en nombre réduit) montrent que la forme n'est pas régulière : elle n'est ni un cercle, ni une ellipse.
- les points qui "s'estompent" puis réapparaissent, éventuellement, montrent l'existence de portions concaves ou convexes. De taille plus ou moins importante, ces structures permettent d'appréhender la taille relative des détails du contour.

La carte CPI permet donc de lire le contenu fréquentiel d'un contour : pour les petites valeurs de σ , on a les détails de hautes fréquences et pour les grandes valeurs de σ , on a les grosses structures qui représentent les basses fréquences. En d'autres termes, lorsque σ tend vers 0, la réponse impulsionnelle du filtre gaussien tend vers une distribution de Dirac qui correspond à un filtre passe-tout expliquant, ainsi, le fait que tous les points du contour sont des points CPI pour $\sigma=0$. Lorsque σ tend vers $+\infty$, la réponse impulsionnelle du filtre gaussien tend vers une constante qui revient à ne garder que la composante continue du signal d'entrée : ce qui explique donc que le contour filtré tende vers le centre de gravité du contour.

En observant une carte CPI (comme celle de la Fig.VII.18), on peut remarquer que pour les grandes valeurs de σ , les variations ayant été éliminées par filtrage, les points d'intersection sont de moins en moins nombreux : leur nombre tend bien vers une limite.

VII.3 - Application à la description des formes

VII.3.1 - Caractérisation de contours

Les figures VII.19 présentent d'autres exemples de contours ainsi que les cartes CPI correspondantes. L'analyse de ces cartes ainsi que celles présentées plus haut montrent clairement que la topologie des cartes CPI diffère fortement d'une forme à l'autre et que par conséquent la fonction CPI est un bon candidat pour la définition des attributs caractéristiques dans le contexte de la reconnaissance des formes. Ainsi, les lettres "M" et "W" (extraites du jeu de caractères *bold.chr* de Borland®) qui présentent une certaine ressemblance (à une rotation 180° près) produisent des cartes CPI totalement différentes donc des attributs issus de la carte CPI permettront de discriminer ces deux lettres.

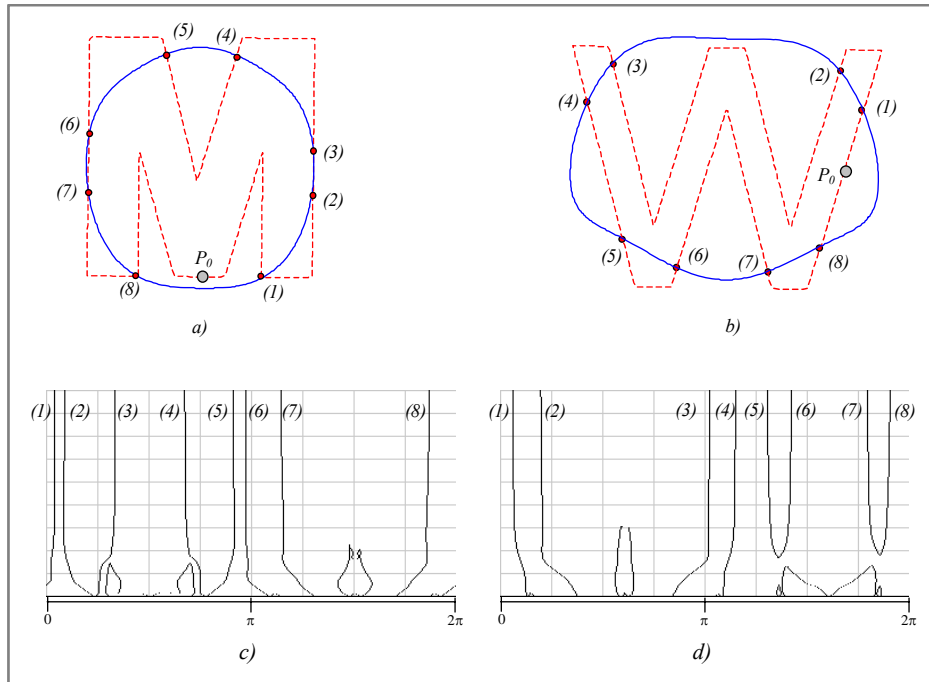


Fig.VII.19 : Contours et cartes CPI des lettres "M" et "W" extraites du jeu de caractères bold.chr de Borland®

VII.3.2 - La carte CPI et le changement d'échelle

Comme on peut le voir sur les figures VII.20, la fonction CPI reste insensible au changement d'échelle pour un facteur de réduction d'échelle supérieur à un seuil qui est lié à la forme et à la taille. Il est clair que l'opération de réduction qui correspond à un sous-échantillonnage conduit à une altération totale ou partielle de la forme si le facteur de réduction devient trop faible.

Un agrandissement, quant à lui, reste sans effet sur le traitement, car cette opération qui correspond à un sur-échantillonnage est compensée lors du prétraitement de normalisation.

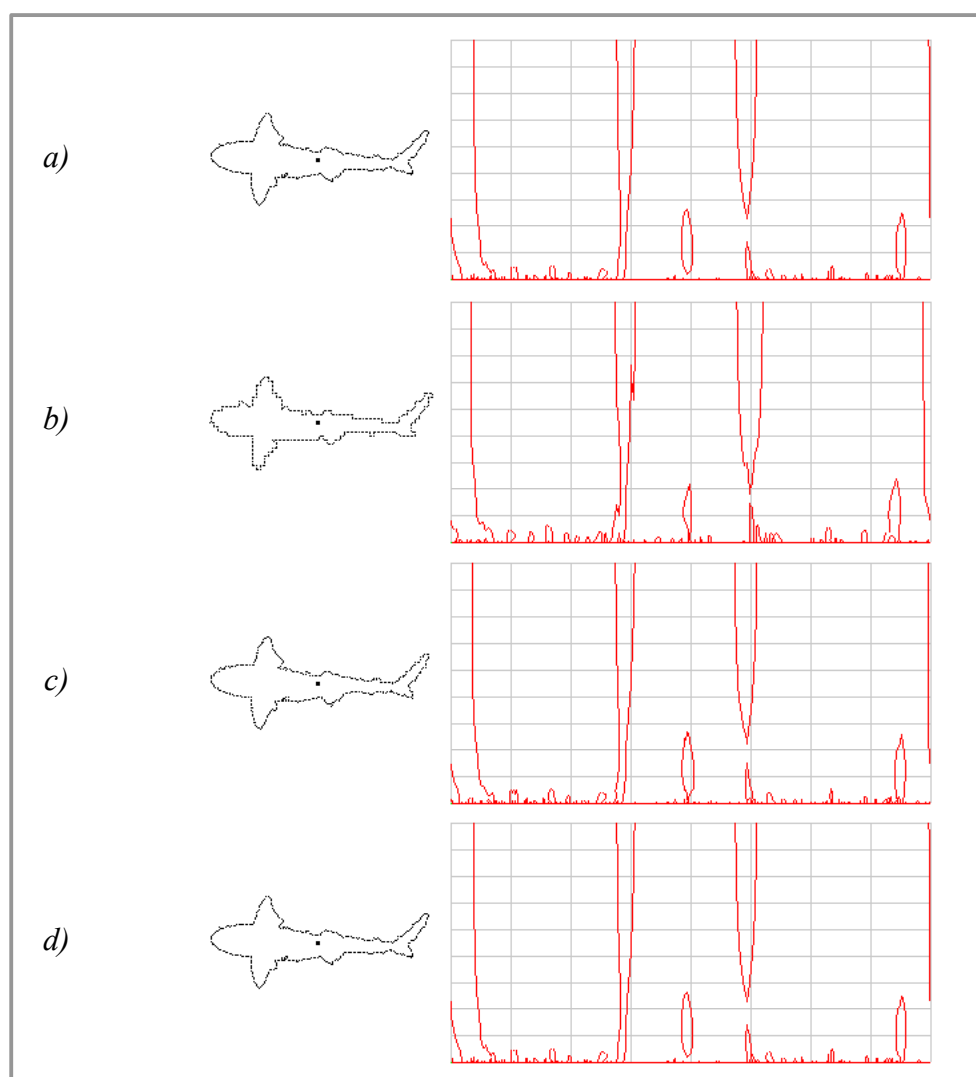


Fig.VII.20 : Effet du changement d'échelle sur la carte CPI

Cette figure montre les cartes CPI correspondant à différents facteurs de changement d'échelle. Sur la figure VII.20.a est représentée la forme originale (sans changement d'échelle). Les figures VII.20.b-d montrent les CPI de la même forme après un changement d'échelle de facteur $\times 0.15$, $\times 0.5$ et $\times 1.5$. A part la forme de la VII.20.b qui donne une carte CPI légèrement altérée les autres donnent une CPI identique à celle de la forme originale même celle avec une réduction d'échelle de 0.5.

Il faut noter que le changement d'échelle est réalisé en se plaçant dans un contexte de traitement d'image (représentation *bitmap* par opposition à une représentation vectorielle) et donc que les coordonnées des contours sont arrondies en valeurs entières afin qu'elles restent compatibles avec les pixels dans une image. Ceci explique la détérioration de la carte CPI qui serait préservée si le traitement était effectué dans un contexte vectoriel où il n'y aurait pas besoin d'arrondir les coordonnées. Nous avons, néanmoins, choisi de rester dans le contexte *bitmap* qui est donc le moins favorable sachant qu'avec une représentation vectorielle, les résultats ne seraient que meilleurs.

VII.3.3 - La carte CPI et la rotation

Comme le montrent les figures VII.21, la carte CPI est insensible à la rotation. La topologie de la carte CPI reste identique quel que soit l'angle de rotation à l'exception près d'une translation circulaire qui permet de déterminer cet angle de rotation. Sur la figure VII.21.a), on représente la carte CPI d'une forme. Les autres figures (VII.21.b-d) représentent les cartes CPI de la même forme après une rotation de 45° , 90° et 180° respectivement. Il apparaît clairement sur ces figures que la carte CPI d'une même forme reste topologiquement invariable.

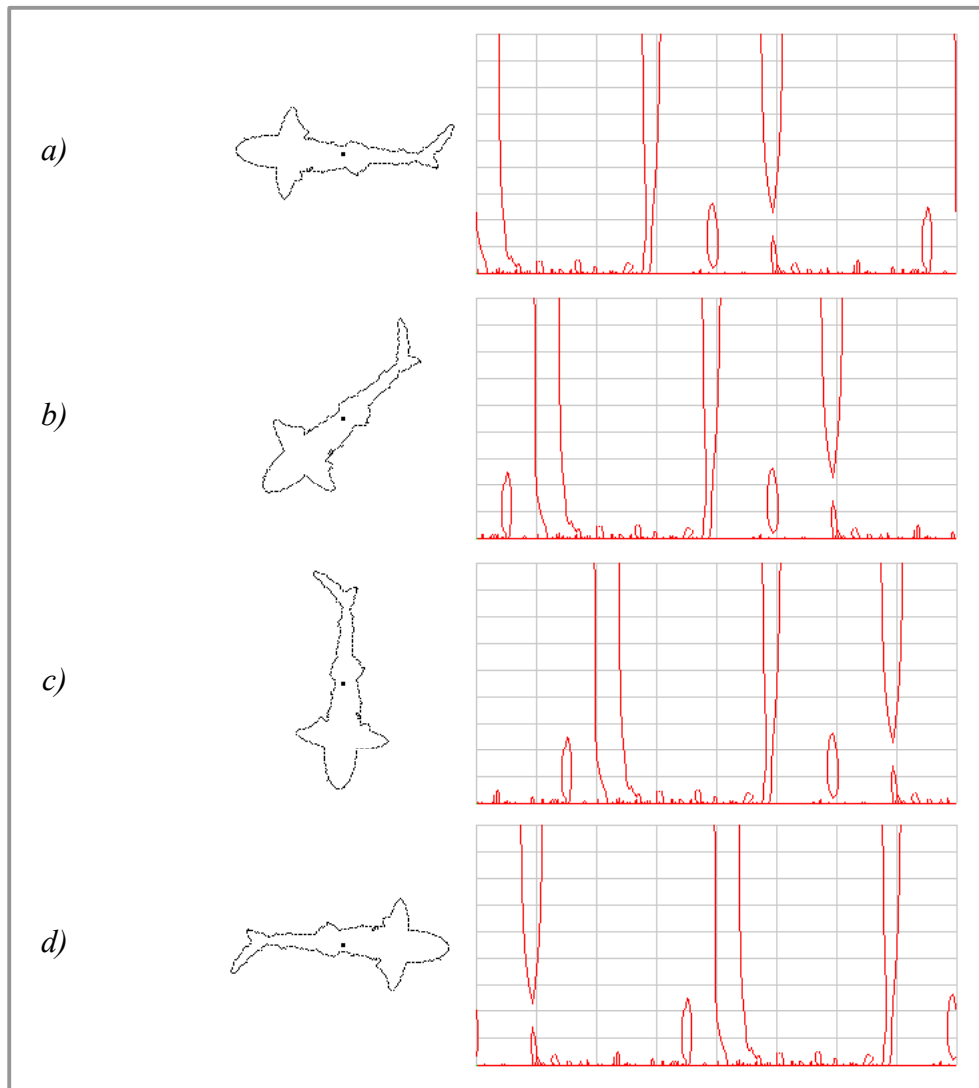


Fig.VII.21 : Effet de la rotation sur la carte CPI

Une propriété intéressante de cette méthode est la lecture directe de l'angle de rotation sur la carte CPI. En effet, si l'on considère les cartes CPI de deux contours dont l'un correspond à une rotation de l'autre, il suffit de repérer la "trace" CPI_0 d'un point d'intersection caractéristique sur les deux cartes et de prendre la différence de leurs abscisses pour obtenir l'angle de rotation θ de l'un des contours par rapport à l'autre. La figure VII.22 illustre bien cette lecture directe de l'angle de rotation : en comparant deux à deux les cartes CPI des trois formes, on peut lire directement l'angle de rotation :

- de la forme VII.22.b par rapport à VII.22.a qui vaut $\theta_0=45^\circ$,
- de la forme VII.22.c par rapport à VII.22.b qui vaut $\theta_1=45^\circ$,

- de la forme VII.22.c par rapport à VII.22.a qui vaut $\theta_2=90^\circ$,

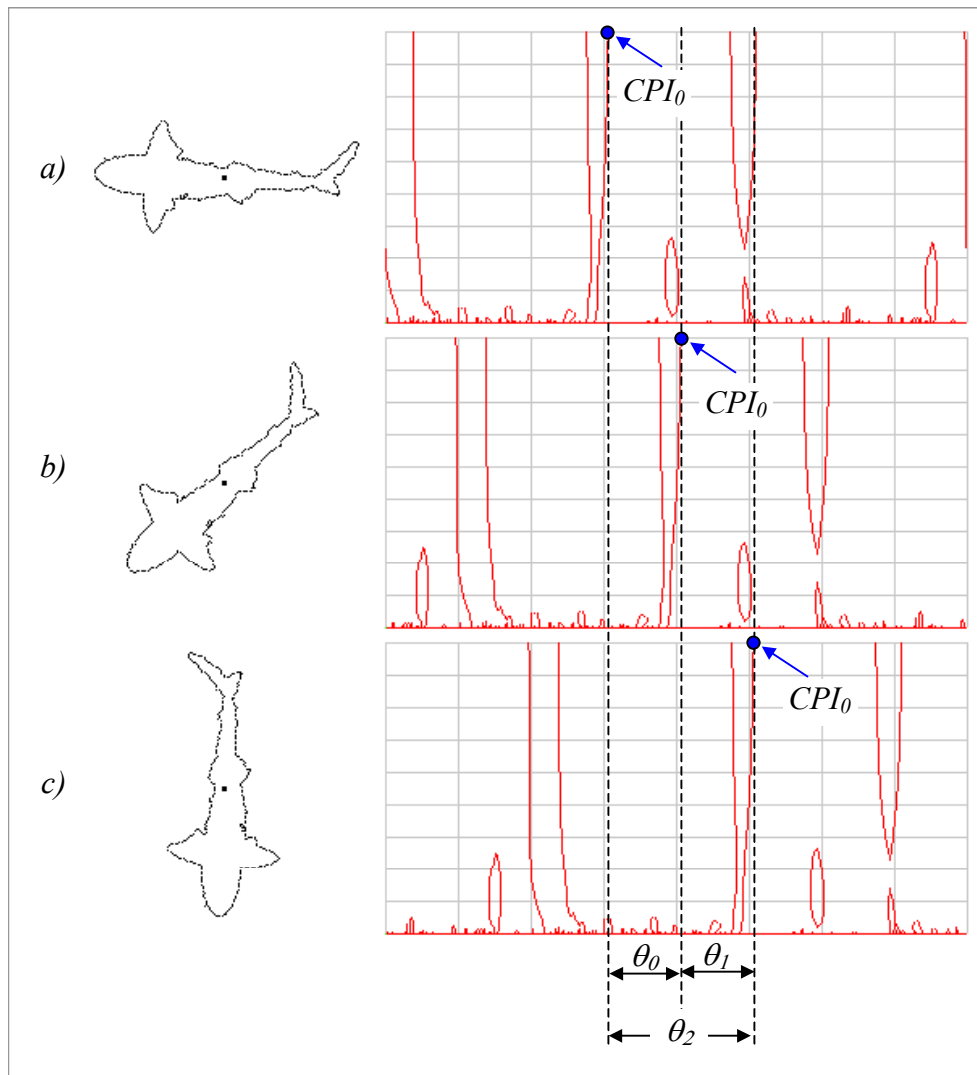


Fig.VII.22 : Détermination de l'angle de rotation par lecture directe

VII.3.4 - Application à l'analyse des détails

Les structures représentant les portions convexes/concaves, plus précisément les détails de composantes de hautes fréquences du contour peuvent utilement être analysées pour la détection des défauts sur une forme. En effet, la forme test F étant reconnue pour être le modèle F_0 de la base de données, on pourra alors réaliser une analyse plus fine recourant à de faibles valeurs de σ afin d'évaluer la "qualité" de F par rapport au modèle F_0 . En réduisant le pas de progression de σ , on obtient une carte CPI plus précise révélant ainsi les plus fins détails du contour. La figure VII.23 présente trois cartes CPI avec des pas de progression d'échantillonnage différents pour σ . En passant d'un pas d'échantillonnage de 1 en 23.a à 0.5 en 23.b puis à 0.25 en 23.c, on rehausse les petites structures permettant ainsi d'appréhender plus facilement le contenu du contour dans les hautes fréquences. Sur la carte CPI de cette figure, le cercle (ou le rectangle) indique la représentation de la portion correspondante du contour entourée respectivement par un cercle (ou un rectangle) sur la forme. On voit bien que les fins détails ressortent d'autant mieux, dans la carte CPI, que l'analyse est plus fine.

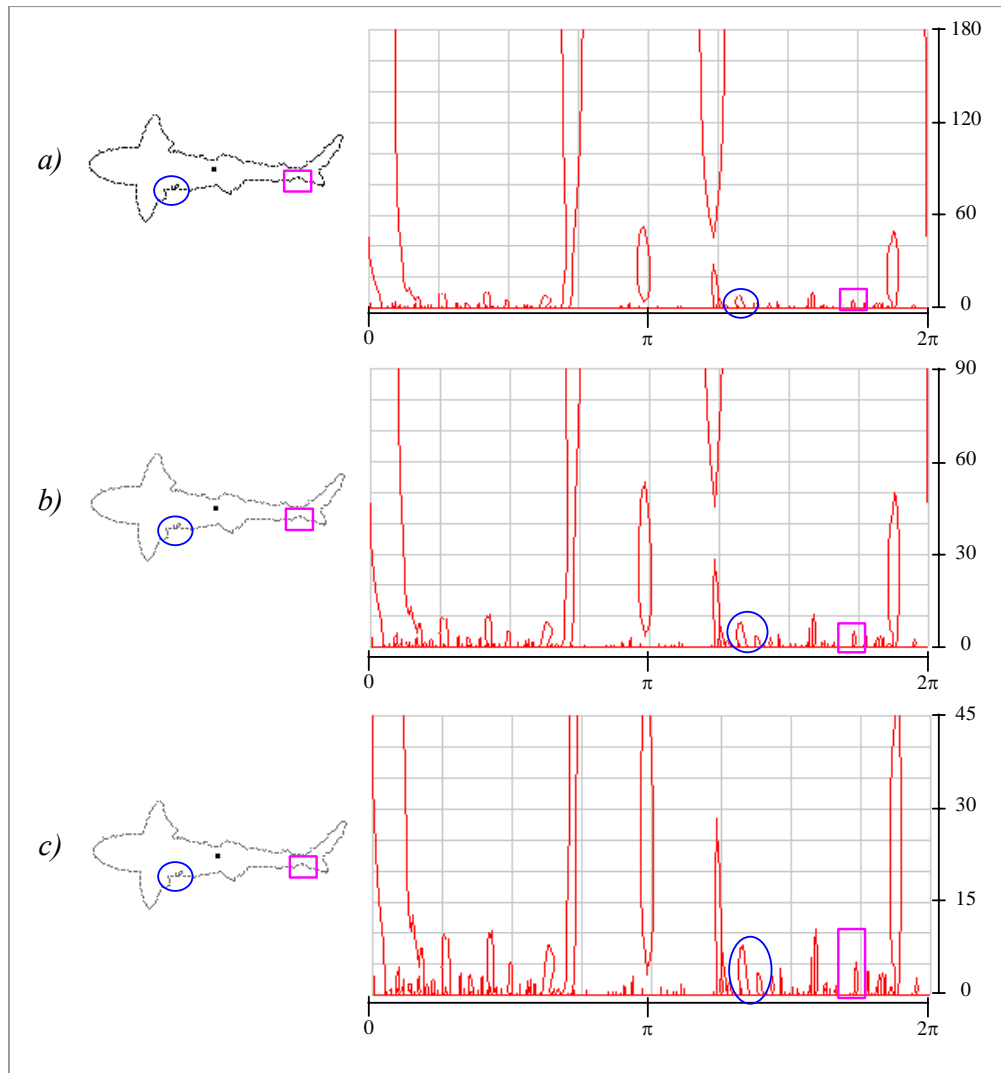


Fig.VII.23 : La carte CPI des détails

VII.3.5 - La carte CPI et le bruit

La soumission de cette approche à des contours bruités révèle une certaine résistance au bruit. Sur la figure VII.24, où la forme se voit perturbée par un bruit, nous pouvons noter les points suivants : globalement la carte CPI n'est pas trop affectée par le bruit surtout aux grandes échelles (voir la partie supérieure de la carte CPI) ; l'apparition de nombreuses petites structures dans la partie basse de la carte CPI. Ce dernier point indique que l'analyse subit l'effet du bruit essentiellement dans les hautes fréquences. Le bruit a introduit des détails fins (de hautes fréquences), expliquant l'apparition de ces petites structures dans les petites échelles de la carte CPI.

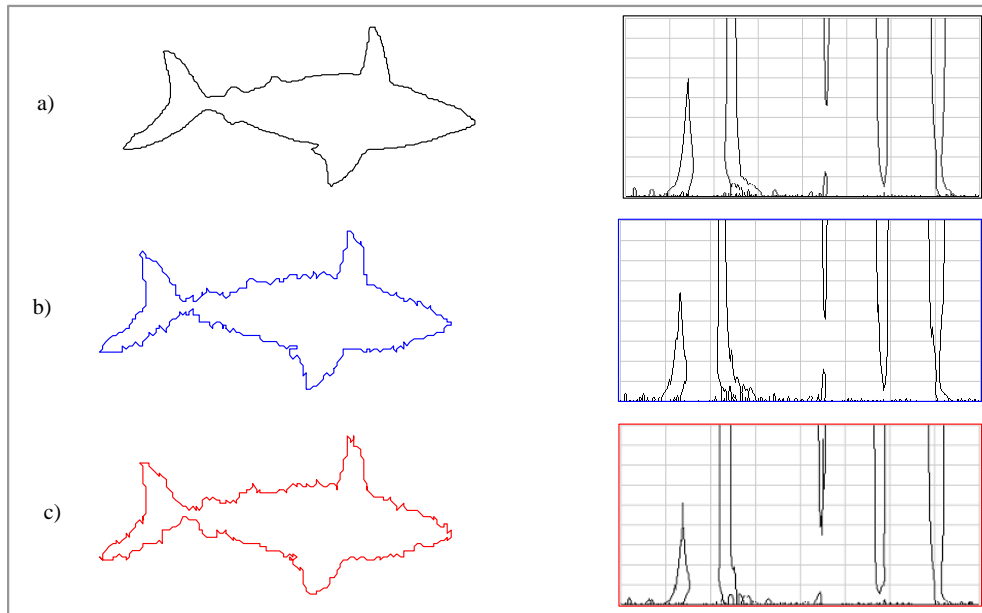


Fig.VII.24 : Effet du bruit sur la carte CPI

VII.3.6 - Définition des attributs

En analysant l'allure de la carte CPI, on peut noter que certaines mesures restent invariantes quelle que soit la transformation. Ces mesures peuvent donc servir d'attributs pour caractériser le contour d'une forme plane dans le cadre d'un processus de reconnaissance des formes. Dans l'état actuel de l'étude, nous pouvons proposer les mesures suivantes comme attributs caractéristiques :

- le nombre limite N_L des points d'intersection ($\sigma=180$),
- la distance (circulaire*) minimale entre deux points consécutifs p_i et p_{i+1} des N_L points limites,
- la distance entre deux points consécutifs. En déterminant les points d'intersection les plus éloignés p_i et p_{i+1} , on peut déterminer les distances entre deux points consécutifs et définir les $d(p_i, p_{i+1})$ comme attributs et cela pour l'ensemble des N_L points d'intersection en prenant le point p_i comme point de départ.
- les mesures caractéristiques des parties concaves et convexes du contour. En effet, comme il a été dit plus haut, pour les contours irréguliers, la carte CPI comporte des structures indiquant la localisation de ces portions convexes et/ou concaves : la hauteur, la surface et la largeur (à une valeur donnée de σ) de ces structures peuvent aussi constituer des attributs caractéristiques d'un contour.
- les distances entre les points CPI à une valeur donnée de σ . En fixant σ à la valeur σ_0 , on dispose de $N(\sigma_0)$ points CPI. En déterminant, comme précédemment, les deux points consécutifs les plus éloignés, p_a et p_b , on définit un vecteur attribut dont les composantes sont les distances circulaires entre les points CPI consécutifs en partant de p_a .

* Le paramètre u étant normalisé à 2π , la notion de distance circulaire veut que si le point p_i est d'abscisse $u_i < 2\pi$ et que le point p_{i+1} est d'abscisse $u_{i+1} < u_i$ alors $d(p_i, p_{i+1}) = (2\pi - u_i) + u_{i+1}$.

VII.3.7 - Application à la reconnaissance des formes

VII.3.7.1 - Choix des attributs

La figure VII.25.a montre la carte CPI d'une forme. Il apparaît clairement que le nombre de points CPI, et par conséquent, la distance entre deux points consécutifs n'est pas la même en fonction de la valeur de σ . Alors que nous avons 6 points CPI en VII.25.a, nous en avons 8 en VII.25.b.

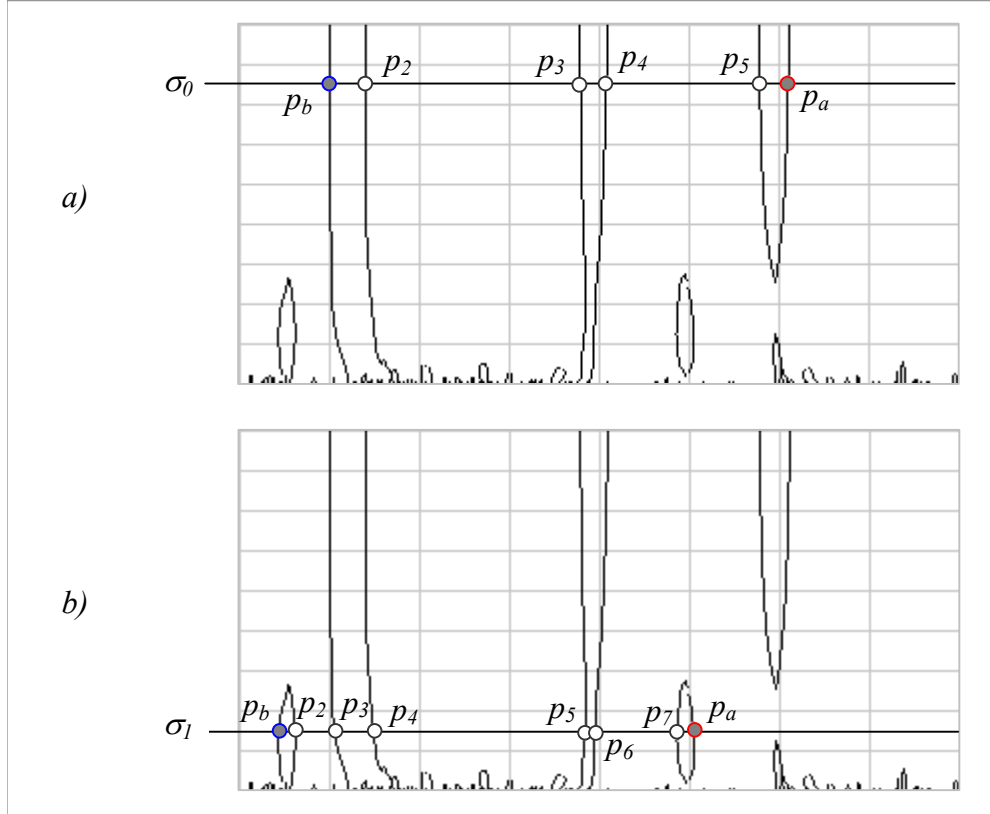


Fig.VII.25 : Définition des attributs en fixant la valeur de σ .

Comme il a été dit plus haut, si pour une valeur σ_0 de σ on détermine les points p_a et p_b , les points consécutifs les plus éloignés sur la carte CPI, on peut indexer l'ensemble des points CPI de telle sorte que $p_0 = p_a$, $p_1 = p_b$, ..., $p_N = p_a$; les points intermédiaires, quant à eux suivent naturellement leur prédécesseur sur la carte.

Une fois que cela est fait, il ne reste plus qu'à définir un vecteur attribut $w(\sigma_0)$ à $N(\sigma_0)$ dimensions dont les composantes sont les distances circulaires entre deux points consécutifs.

$$w(\sigma_0) = (d_{0,1} \quad d_{1,2} \quad \dots \quad d_{i,i+1} \quad \dots \quad d_{p_{N(\sigma_0)-1}, p_{N(\sigma_0)}})^T \quad (15)$$

La distance circulaire $d_{i,i+1}$ entre deux points consécutifs d'abscisses respectives u_i et u_{i+1} étant définie par :

$$d_{i,i+1} = d(p_i, p_{i+1}) = \begin{cases} u_{i+1} - u_i & \text{si } u_{i+1} \geq u_i \\ 2\pi + u_{i+1} - u_i & \text{si non} \end{cases} \quad (16)$$

Le vecteur $w(\sigma_0)$ étant, généralement, différent pour chaque valeur de σ_0 , il est envisageable de fixer un certain nombre de valeurs de σ et de définir un vecteur attribut V global de dimension plus importante. Ainsi, si K est le nombre de valeurs de σ fixées, on obtient un vecteur V tel que :

$$V = (w(\sigma_0) \quad w(\sigma_1) \quad \dots \quad w(\sigma_k) \quad \dots \quad w(\sigma_{K-1}))^T \quad (17)$$

Le vecteur V est alors de dimension

$$N = \sum_{k=0}^{K-1} N(\sigma_k) \quad (18)$$

Ainsi, on prend en compte de l'information issue de plusieurs échelles et de ce fait, le vecteur V devient un descripteur multiéchelle.

VII.3.7.2 - Mesure de similarité

Après la définition des attributs, l'étape suivante, et non la moindre, est la mesure de similarité qui donne le degré (ou le taux) de ressemblance entre deux objets A et B. Dans une telle situation, la fonction la plus immédiate est la mesure de distance entre les vecteurs attributs V_A et V_B dans l'espace à N dimensions dans une norme définie au préalable et cela en fonction de la nature des vecteurs attributs (norme euclidienne, Manhattan, de Minkowski d'ordre p, ...).

Pour notre étude, nous avons choisi d'utiliser une fonction basée sur le cosinus de l'angle γ entre les vecteurs V_A et V_B . Etant donné deux vecteurs V_A et V_B dans un espace à N dimensions, on a :

$$\cos(\gamma) = \frac{\sum_{i=0}^{N-1} V_A^i \cdot V_B^i}{\sqrt{\sum_{i=0}^{N-1} (V_A^i)^2 \cdot \sum_{i=0}^{N-1} (V_B^i)^2}} \quad (19)$$

qui varie entre -1 et 1 . On peut alors définir une fonction de similarité S par

$$S(A, B) = S(\gamma) = 50(1 + \cos(\gamma)) \quad (20)$$

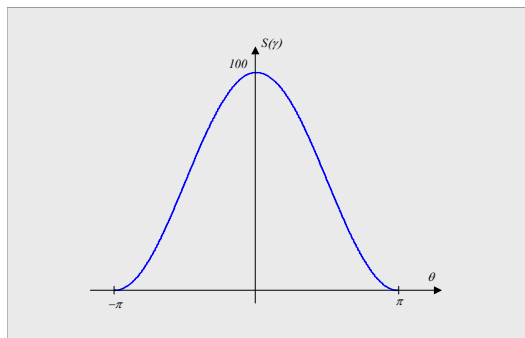


Fig.VII.25 : Allure de la fonction de similarité S

Par définition, cette fonction donne une similarité de 100% lorsque les deux objets sont identiques (les deux vecteurs sont en phase) et 0% lorsque l'angle entre leurs vecteurs attributs est de π : les deux vecteurs sont alors en opposition de phase.

Pour ce qui concerne notre méthode, l'inconvénient avec cette fonction est qu'il existe une ambiguïté qu'il faudra lever avant la décision. En effet, pour deux vecteurs en phase, la fonction S donne un taux de 100% alors qu'ils peuvent être très différents en module : c'est le cas des vecteurs U , V et W de la figure VII.26.

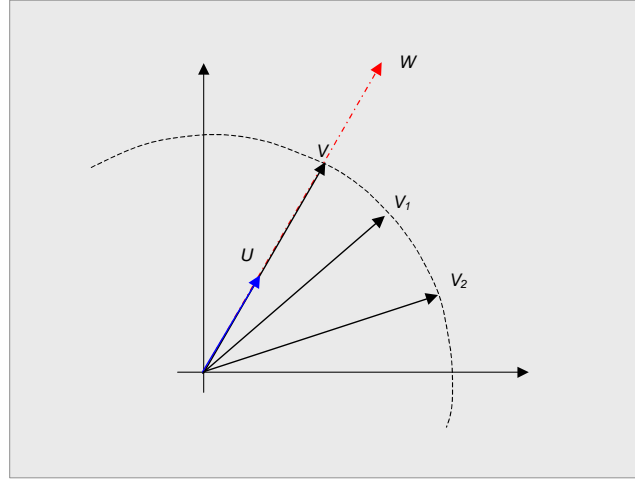


Fig.VII.26 : Exemples de vecteurs dans un espace

Pour lever cette ambiguïté, nous proposons d'utiliser la fonction c suivante qui introduit un coût dans la fonction de similarité :

$$c(\delta) = c(A, B) = \delta = \frac{\text{Min}(\|V_A\|, \|V_B\|)}{\text{Max}(\|V_A\|, \|V_B\|)} \quad (21)$$

où $\|V_A\|$ est le module du vecteur V_A . Cette fonction varie entre 0 et 1 : elle vaut 1 si les deux vecteurs sont de module identique et 0 lorsque l'un est nul. Prise indépendamment, la fonction c engendre aussi une ambiguïté pour tous les vecteurs de même module comme c'est le cas pour les vecteurs V , V_1 et V_2 de la figure VII.26.

En combinant les deux fonctions, on définit une nouvelle fonction similarité *SimScore* par :

$$\text{SimScore}(A, B) = \text{SimScore}(\delta, \gamma) = S(A, B) \cdot c(A, B) = 50 \cdot \delta \cdot (1 + \cos(\gamma))^{(*)} \quad (22)$$

on obtient une mesure de 100% lorsque les deux vecteurs sont égaux et 0% lorsqu'ils sont en opposition de phase ou lorsque l'un est nul. La figure VII.27 montre cette fonction de mesure de similarité, sous forme d'image en VII.27.a et en 3D en VII.27.b. Comme on peut le voir sur ces figures, la fonction $\text{SimScore}(\delta, \gamma)$ atteint son maximum en $\delta=1$ et $\gamma=0^\circ$ qui correspondent à une égalité parfaite entre les vecteurs à comparer. Sur ces figures, la similarité varie du blanc (100% de similarité) au noir (différence "total") en passant par le gris qui donne le taux de ressemblance intermédiaire entre deux vecteurs.

(*) Pour la suite, on utilisera indifféremment $\text{SimScore}(A, B)$ ou $\text{SimScore}(\delta, \gamma)$ pour indiquer la fonction de similarité entre les formes A et B : δ étant le rapport $\frac{\text{Min}(\|V_A\|, \|V_B\|)}{\text{Max}(\|V_A\|, \|V_B\|)}$ et γ l'angle entre les deux vecteurs attributs V_A et V_B .

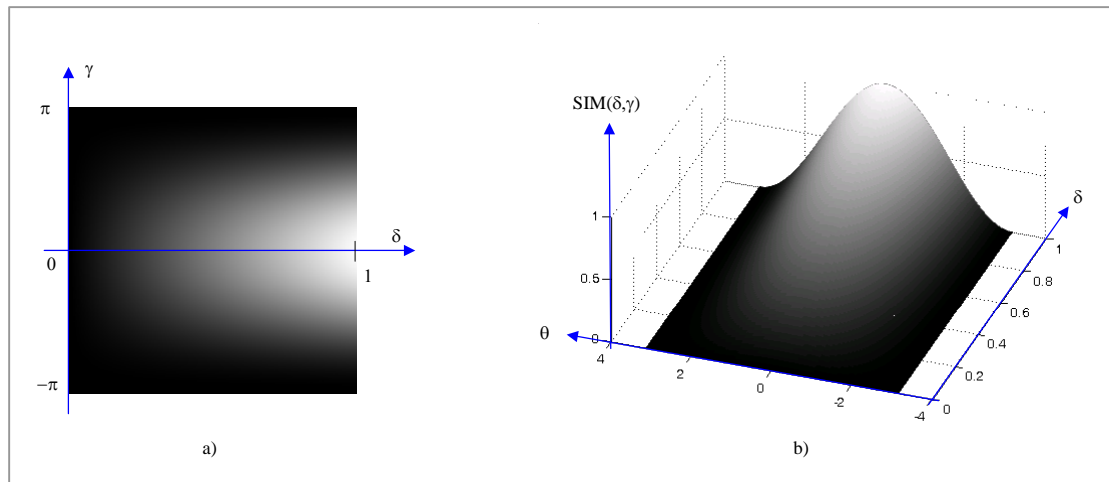


Fig.VII.27 : La fonction de similarité SimScore

VII.3.7.3 - Exemples d'application à la reconnaissance

VII.3.7.3.1 - Description du contexte expérimental

Pour cette première étude d'évaluation, nous disposons d'une base de 1292 objets de diverses formes représentées par leurs contours. On dispose ainsi de formes purement synthétiques, de contours des images d'animaux marins extraits de la base d'images CVSSP^(*) [1] et de caractères alphanumériques typographiques (extraits du jeu de caractères *bold.chr* de Borland[®]). Les figures VII.28 suivantes montrent un échantillon de ces formes.

Dans l'état actuel de l'étude, lorsque la forme comporte plusieurs contours comme c'est le cas pour le chiffre '8' ou pour la lettre 'B', nous ne prenons en compte que les contours les plus externes.

^(*) Centre for Vision, Speech, and Signal Processing (CVSSP) avait mis une base de contours téléchargeable à la disposition des chercheurs, <http://www.ee.surrey.ac.uk/Research/VSSP/imagedb> : (lien brisé depuis quelque temps!).

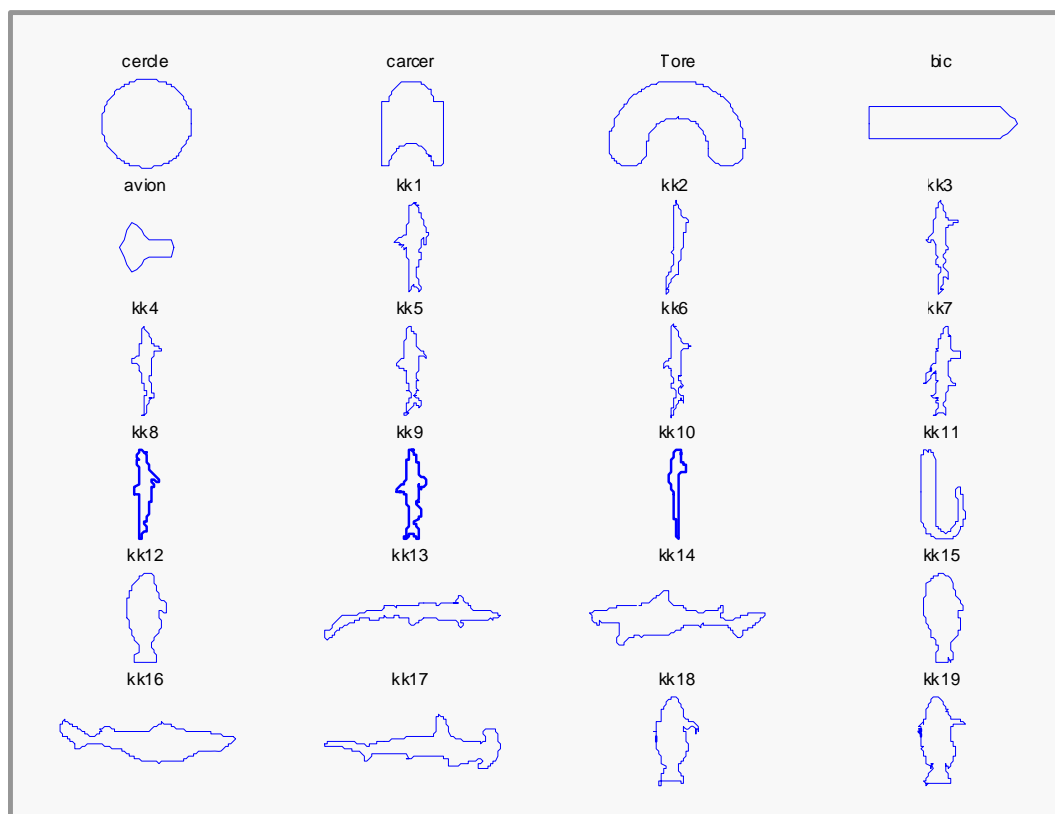


Fig.VII.28.a : Un échantillon de formes

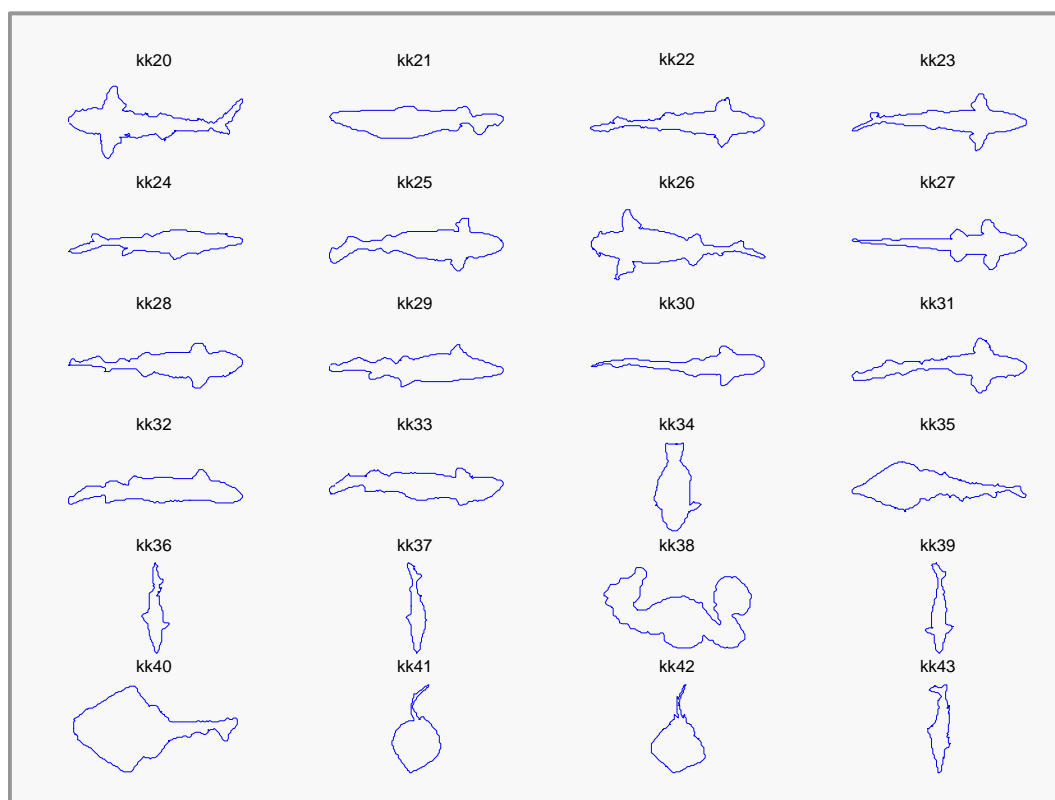


Fig.VII.28.b : Un échantillon de formes (suite)

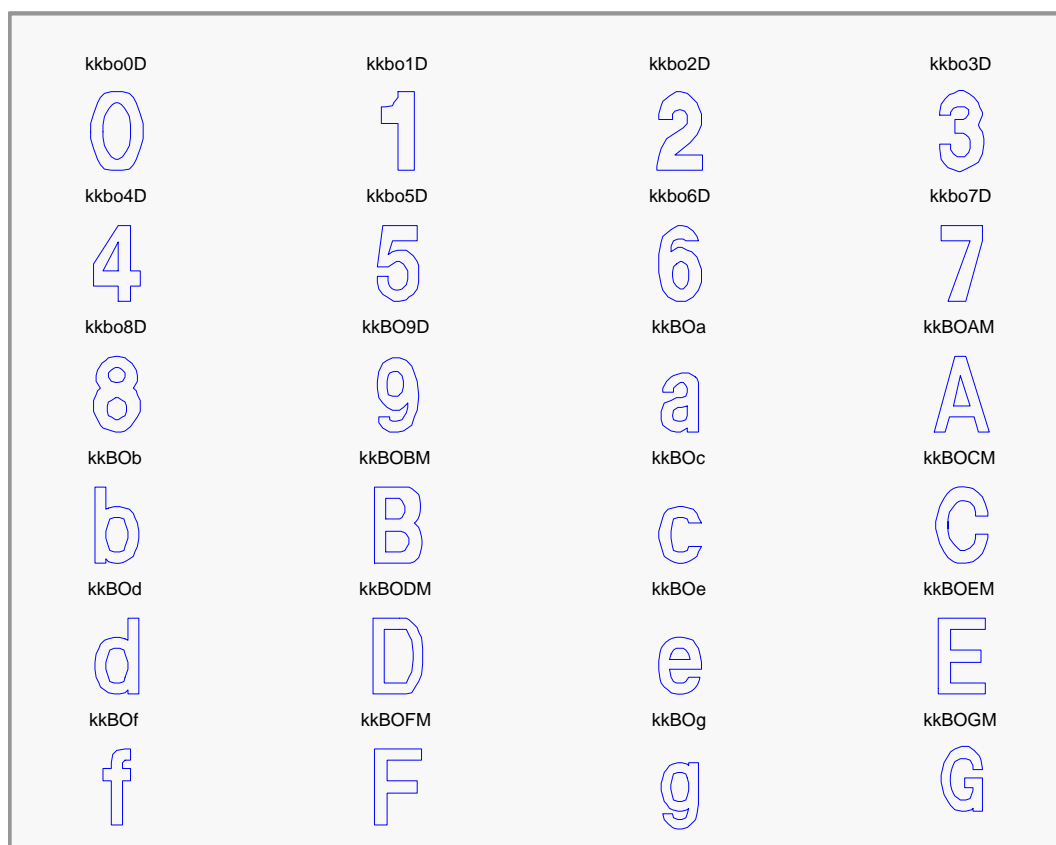


Fig.VII.28.c : Un échantillon de formes (fin)

Sur les 1292 formes nous avons 1167 formes différentes et 25 qui sont des copies déformées en leur appliquant une rotation et une homothétie. Les 100 formes restantes sont des copies déformées comme précédemment en appliquant une rotation et/ou une homothétie à une même forme. Le tableau suivant donne les différents angles de rotation et les facteurs d'échelle appliqués à certaines formes pour enrichir la base de données dans le cadre de cette étude.

Paramètres	Valeurs des paramètres de transformation									
Angle de rotation	10°	45°	60°	90°	135°	180°	225°	270°	300°	330°
Facteur d'échelle	0.2	0.25	0.5	0.75	1	2	25	10	20	30
Indice n	0	1	2	3	4	5	6	7	8	9

Tableau VII.1 : Paramètres de transformations appliquées

Dans ce contexte, pour faciliter la désignation des formes résultantes, l'application de la rotation d'indice i et du facteur d'échelle (ou facteur d'homothétie) d'indice j à la forme de nom **abc** produit la forme nommée **RiHjabc** comme on le verra plus loin dans cette section. La forme **R1H6kk707** de la figure suivante est donc obtenue en appliquant une rotation de 45° et un facteur d'échelle de 25 à la forme **kk707**.

Nous avons réalisé des expérimentations en utilisant 5 valeurs de σ (180, 20, 30, 40 et 50) pour construire le vecteur attribut comme indiqué par les équations (11-14).

L'application de la reconnaissance consiste à tirer, au hasard, un objet de la base : c'est la forme requête. Le test consiste alors à extraire de cette forme requête les attributs définis préalablement et de comparer son vecteur attribut V_0 aux vecteurs attributs

$V_{i(i=1,2, \dots, 1292)}$ correspondants des 1292 formes de la base en utilisant la fonction de similarité *SimScore* définie par l'équation (18).

La stratégie de reconnaissance se fait en deux étapes :

- la première étape consiste à ne retenir de la base que les objets dont la taille du vecteur attribut est la même que celle de la forme requête,
- à la deuxième étape, on applique la fonction de similarité *SimScore* entre l'objet test et les N retenus dans la première étape. Le résultat du traitement donne ensuite les 8 objets les plus proches de la forme requête dans l'ordre décroissant du taux de similarité.

VII.3.7.3.2 - Forme requête : *kk707*

Dans cet exemple de la figure VII.29, la forme, *Testkk707*, à analyser est une copie, sans aucune déformation, de la forme *kk707*.

Avec les cinq valeurs de σ_0 : 72 formes ont été retenues à la première étape, dont 32 de la famille *kk707* qui ont été récupérées à la deuxième étape avec un score de similarité de 100%.

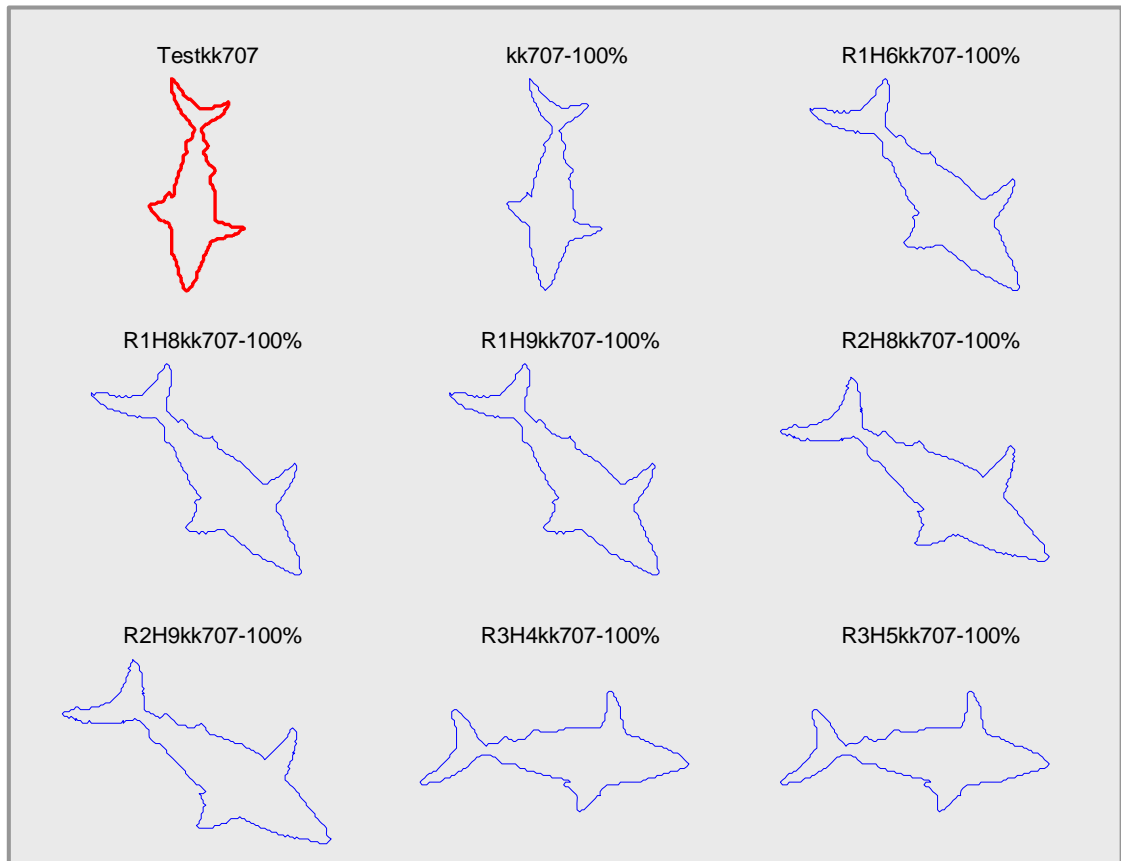


Fig.VII.29 : Résultat de la reconnaissance de la forme *Testkk707* dans la base.

Le tableau VII.2 donne la liste des 72 formes retenues et qui sont reconnues, par la suite, comme étant la forme *Testkk707* avec un taux de similarité plus ou moins élevé et précisé ici par le nombre entre parenthèses.

Ce tableau et la figure VII.29 nous révèlent que l'approche caractérise assez bien les formes puisque les meilleurs taux de reconnaissance sont obtenus avec les formes de la

même famille que la forme requête. Les 61 premières formes identifiées (au sens de la fonction *SimScore*) sont toutes de la famille *kk707* ; plus précisément, toutes les formes retenues avec un score de 100% ne sont que des *kk707*. Le plus faible taux de reconnaissance obtenu pour une forme de cette famille vaut 2% : ce taux est donné par une forme ayant subi une réduction d'échelle.

<i>kk707</i> (100)	<i>R0H9kk707</i> (100)	<i>R1H6kk707</i> (100)	<i>R1H8kk707</i> (100)
<i>R2H7kk707</i> (100)	<i>R2H8kk707</i> (100)	<i>R2H9kk707</i> (100)	<i>R3H4kk707</i> (100)
<i>R3H5kk707</i> (100)	<i>R3H6kk707</i> (100)	<i>R3H7kk707</i> (100)	<i>R3H8kk707</i> (100)
<i>R3H9kk707</i> (100)	<i>R4H7kk707</i> (100)	<i>R5H4kk707</i> (100)	<i>R5H5kk707</i> (100)
<i>R5H6kk707</i> (100)	<i>R5H7kk707</i> (100)	<i>R5H8kk707</i> (100)	<i>R5H9kk707</i> (100)
<i>R6H6kk707</i> (100)	<i>R6H9kk707</i> (100)	<i>R7H4kk707</i> (100)	<i>R7H5kk707</i> (100)
<i>R7H6kk707</i> (100)	<i>R7H7kk707</i> (100)	<i>R7H8kk707</i> (100)	<i>R7H9kk707</i> (100)
<i>R8H8kk707</i> (100)	<i>R8H9kk707</i> (100)	<i>R9H8kk707</i> (100)	<i>R9H9kk707</i> (100)
<i>R0H6kk707</i> (98)	<i>R1H9kk707</i> (98)	<i>R4H6kk707</i> (98)	<i>R8H7kk707</i> (98)
<i>R9H6kk707</i> (98)	<i>R0H8kk707</i> (98)	<i>R6H7kk707</i> (98)	<i>R6H8kk707</i> (98)
<i>R4H9kk707</i> (97)	<i>R8H6kk707</i> (97)	<i>R2H6kk707</i> (95)	<i>R4H8kk707</i> (95)
<i>R1H7kk707</i> (93)	<i>R2H5kk707</i> (87)	<i>R9H7kk707</i> (87)	<i>R8H5kk707</i> (85)
<i>R7H3kk707</i> (76)	<i>R9H5kk707</i> (74)	<i>R6H2kk707</i> (70)	<i>R4H5kk707</i> (66)
<i>R0H5kk707</i> (62)	<i>R8H4kk707</i> (55)	<i>R6H4kk707</i> (54)	<i>R3H2kk707</i> (52)
<i>R5H2kk707</i> (47)	<i>R2H2kk707</i> (45)	<i>R1H2kk707</i> (29)	<i>R8H1kk707</i> (28)
<i>R5H0kk707</i> (16)	<i>kk1097</i> (5)	<i>R8H0kk707</i> (4)	<i>kk882</i> (4)
<i>kk333</i> (3)	<i>kk850</i> (2)	<i>kk929</i> (2)	<i>kk902</i> (2)
<i>kk829</i> (2)	<i>R6H1kk707</i> (2)	<i>kk511</i> (1)	<i>kk916</i> (1)

Tableau VII.2 : Liste des formes reconnues avec cinq valeurs de σ

VII.3.8 - Etude des performances en reconnaissance des formes

En vue de mieux évaluer notre méthode, de rendre plus accessibles les études comparatives et aussi de faciliter les développements et les évolutions de nos études, nous avons développé une plate-forme logicielle intégrant la méthode adaptée de la méthode CSS et la méthode MSGPR. Dénommée plate-forme Logicielle d'Analyse Multi-Échelle pour la Reconnaissance des Formes (LAMEFor), cette plate-forme comporte, dans son développement actuel, les deux méthodes d'analyse : la CSS et la MSGPR. Dans cette section, nous présentons la plate-forme ainsi que les résultats de trois études comparatives entre les deux méthodes d'analyse intégrées.

VII.3.8.1 - Présentation de LAMEFor

Cette plate-forme de développement est structurée comme l'indique le bloc-diagramme de la figure VII.30. Son fonctionnement consiste à effectuer des analyses pour rechercher un contour dans une base de contours, en procédant en deux étapes :

- 1) tout d'abord, on applique une méthode d'analyse (MSGPR ou CSS) à l'ensemble des contours de la base pour en extraire un vecteur d'attributs par contour : ceci permet alors de construire la base des attributs. A chaque élément de la base d'attributs correspond un contour de la base de contours.
- 2) lorsqu'une requête arrive, on lui applique la même méthode d'analyse pour en extraire le vecteur d'attributs correspondant. Il ne reste plus qu'à appliquer la fonction de similarité *SimScore* à ce vecteur conjointement avec chacun des vecteurs de la base d'attributs et récupérer les contours correspondant aux plus forts taux de similarité.

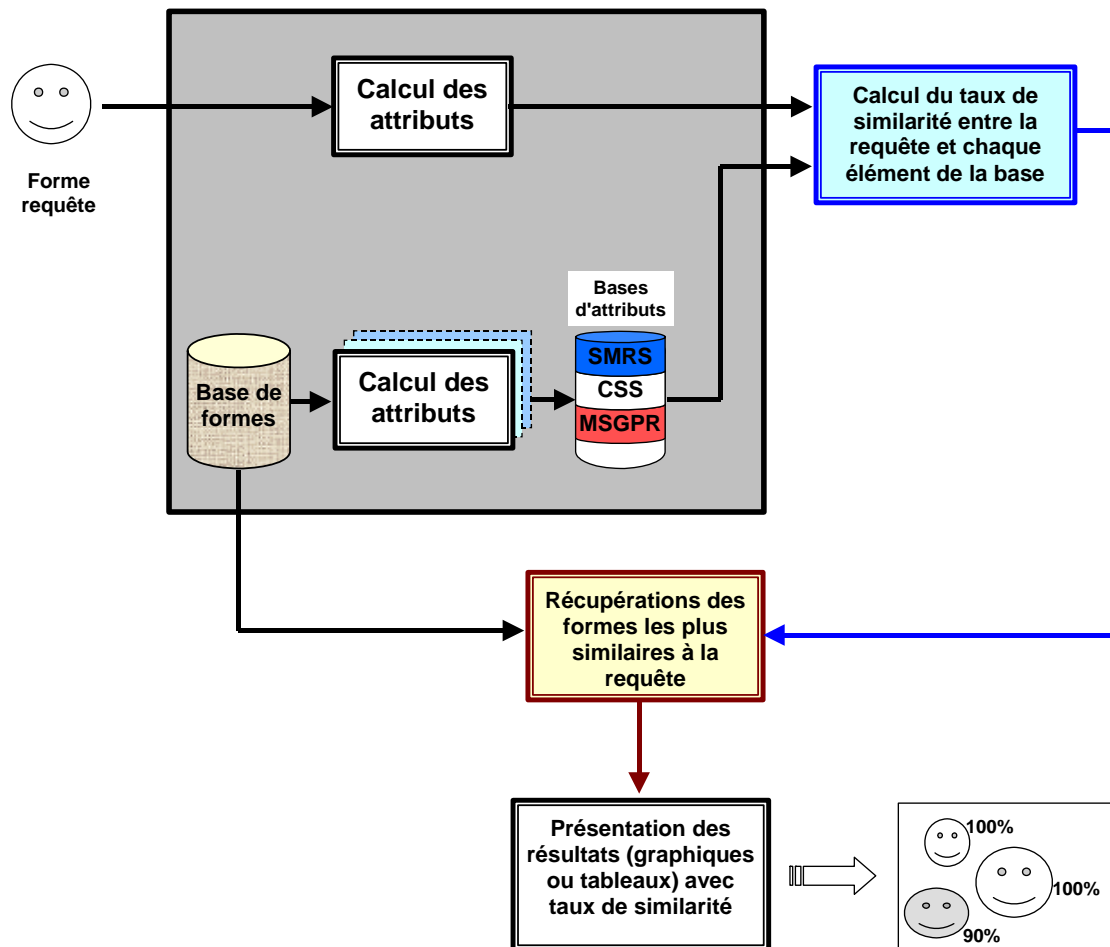


Fig.VII.30 : Diagramme fonctionnel de la plate-forme LAMERFor

VII.3.8.2 - Etude comparative avec la méthode CSS

Pour une meilleure évaluation de notre méthode, nous avons réalisé une étude comparative avec la méthode CSS. Notre étude évalue, ici, le pouvoir discriminant, la résistance aux fausses alarmes et le comportement aux déformations élastiques.

VII.3.8.2.1 - Pouvoir discriminant

Dans une première étude (Fig.VII.31), nous cherchons à évaluer le pouvoir discriminant de l'approche. Il s'agit de mesurer la capacité à récupérer les formes de la famille de la requête existant dans la base avec de forts taux de reconnaissance, d'une part, et récupérer, par contre, les autres formes issues des familles différentes avec de faibles taux, d'autre part. Nous avons choisi comme requête le contour *kk707* pour lequel nous disposons de copies déformées dans la base de données (voir *Description du contexte expérimental*). Avec des attributs dérivés de la carte CPI pour quatre valeurs d'échelle ($\sigma=180, 64, 96, 128$), nous avons pu identifier les formes de la famille avec des scores de similarité allant de 91% à 100%. Le plus haut score fourni par un contour qui ne soit pas de la famille de la requête est de 43%. Avec des attributs issus de l'adaptation de la méthode CSS, les scores des formes de la famille vont de 69% à 100%. Par contre ici, le plus haut score donné par un contour n'appartenant pas à la famille *kk707* est de 68%.

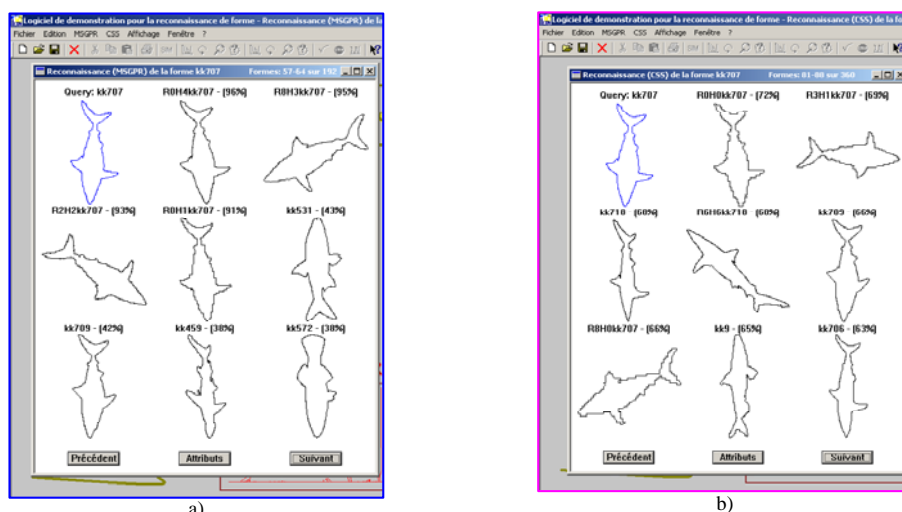


Fig.VII.31 : Recherche dans une base : a) résultats avec les attributs IPM et b) résultats avec les attributs CSS.

Dans cette étude, l'écart entre le plus faible score fourni par un contour de la famille et le plus fort score fourni par un contour de famille différente constitue un indicateur important : il est clair que plus il est fort et plus nette sera la séparation entre les différentes familles.

Avec les attributs de notre approche, le plus bas score de similarité obtenu à partir d'un modèle parent de la requête est de 91% et le plus haut score obtenu à partir d'un modèle qui n'est pas de la famille de la requête est de 43% : la différence entre les deux scores extrêmes est de $\Delta_1=48\%$. Avec la méthode adaptée de CSS, cet écart n'est plus que de $\Delta_2=1\%$ tandis que le plus bas score de similarité obtenu avec les contours parents est de 69%. Naturellement, un grand écart indique une bonne séparation entre les classes : en d'autres termes, la meilleure méthode est celle qui préserve le plus grand écart.

VII.3.8.2.2 - Résistance aux fausses alarmes

Cette expérimentation évalue la résistance de l'approche aux fausses alarmes. Les fausses alarmes sont constituées de l'ensemble de toutes les formes reconnues pour être la requête alors que celle-ci n'existe pas dans la base. Pour cette étude, la requête choisie (*vase*) est bien-sûr inexistante dans la base (Fig.VII.32). Dans cette expérience, plus le score de similarité maximum reste faible et meilleur est le système d'identification : le contour requête n'étant pas dans la base, il vaut mieux que le plus fort score soit le plus faible possible protégeant, ainsi, la méthode des fausses alarmes.

La méthode basée sur la carte CPI récupère 12 modèles avec un score de similarité maximum de 26% tandis que la méthode adaptée de CSS récupère 58 modèles avec un score de similarité maximum de 50%.

Notre méthode donne un score de similarité maximum qui est largement inférieur à celui donné par la méthode adaptée de CSS.

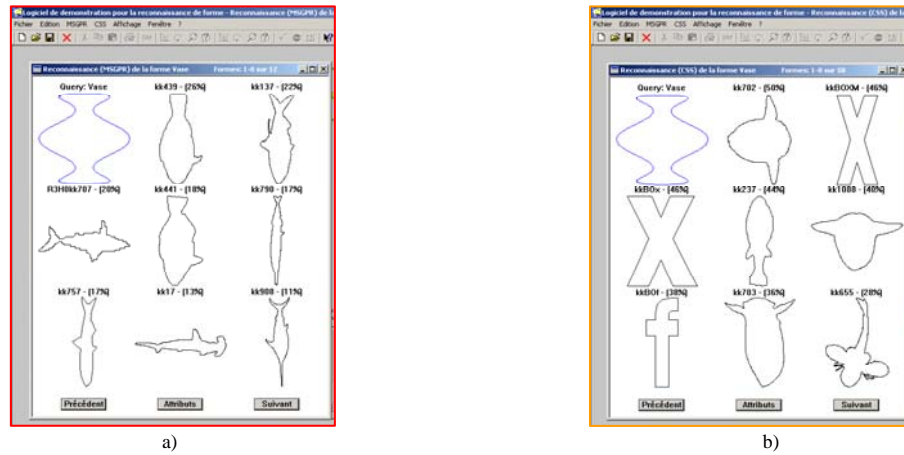


Fig.VII.32 : Recherche d'une forme n'existant pas dans la base : a) résultats avec les attributs IPM et b) résultats avec les attributs CSS.

VII.3.8.2.3 - Comportement aux déformations élastiques

Il s'agit, ici, d'évaluer l'aptitude de notre approche à reconnaître une forme lorsqu'elle est soumise à des déformations élastiques.

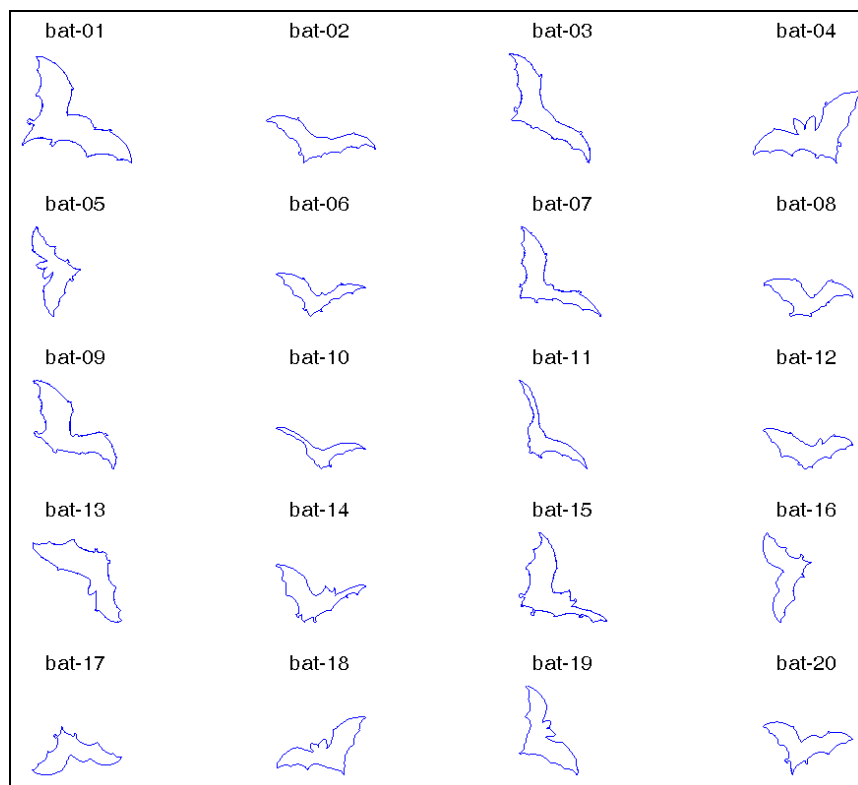


Fig.VII.33 : Exemple de contours de la base expérimentale de MPEG-7

Nous avons appliqué les deux méthodes à un échantillon de la base expérimentale CE-Shape-1 de MPEG-7. Cette expérience consiste à rechercher une requête dans une base de données se composant de vingt contours. Chacun de ces contours est issu du même modèle soumis à différentes transformations rigides ou élastiques. La figure VII.33 montre un exemple de tels contours : il s'agit des contours extraits d'une série d'images d'une chauve-souris (série "bat") en vol prise sous différents angles. Ici, les

déformations sont dues à l'angle de vue et au mouvement : ce sont bien des déformations de type élastique.

Dans cette étude, la méthode adaptée de CSS donne, par contre, de meilleurs résultats que notre méthode. Comme on peut le voir sur la figure VII.33, certains changements d'un contour à l'autre sont des déformations fluides dues au mouvement de l'animal. Comme notre méthode est essentiellement une méthode globale, elle comporte des risques d'échec pour des déformations plus ou moins fluides. Elle est adaptée, en fait, aux déformations rigides telles qu'elles sont définies pour la partie A de l'expérimentation dans [Latecki00, Mokhtarian03]. Dans ces articles, les auteurs ont considéré trois étapes différentes selon les trois parties de l'expérience CE-Shape-1 définies dans le cadre du consortium MPEG-7 : la partie A qui traite de la robustesse au changement d'échelle et à la rotation est considérée comme condition nécessaire dans un processus de reconnaissance des formes. Notre méthode satisfait bien cette condition nécessaire dans la description des formes.

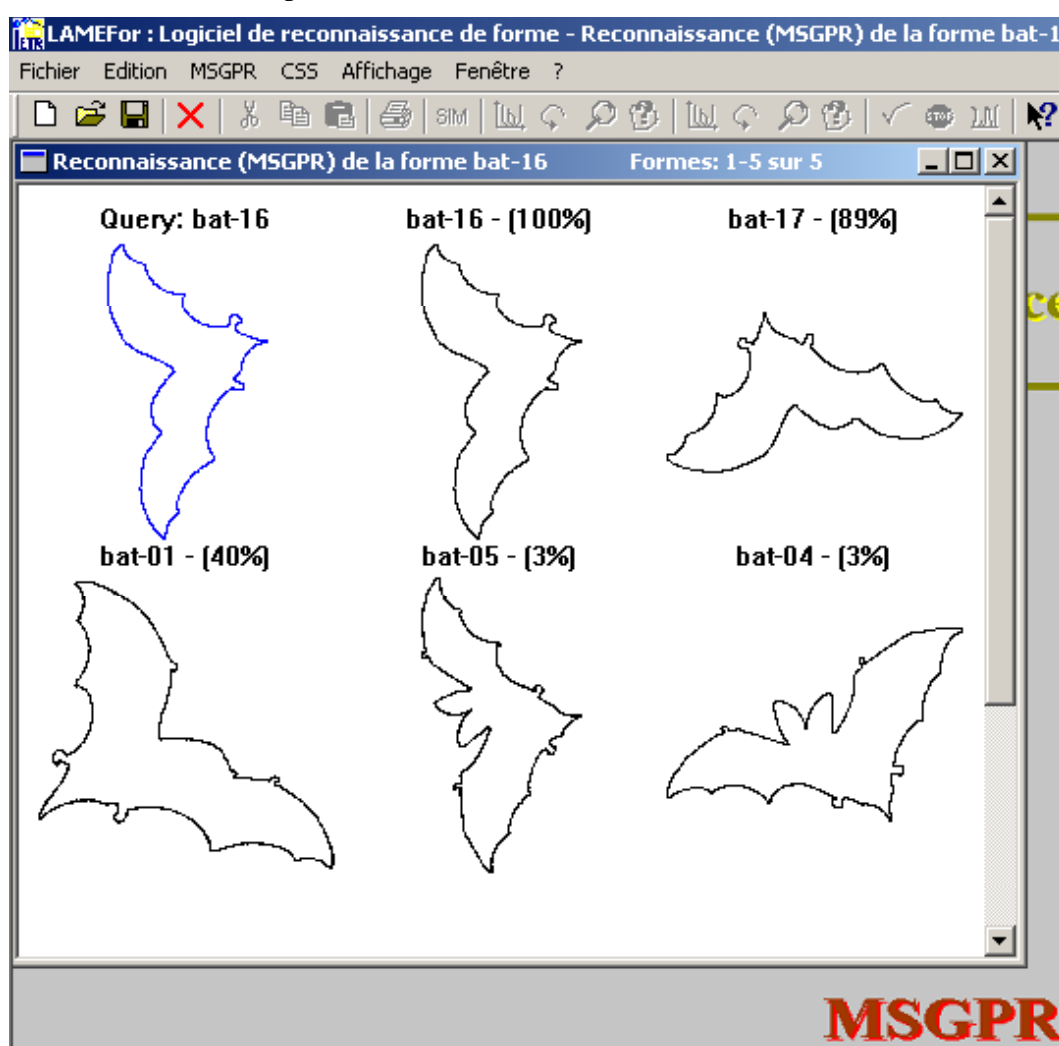


Fig.VII.33 : Résultats de récupération d'une forme dans la base de contours "bats".

Dans tous les cas, la requête, quand elle-même est présente dans la base de données, est récupérée avec un taux de similarité de 100%. Comme on peut le voir sur la figure VII.33, le contour *bat-16*, qui est la requête, est récupéré avec un score de similarité de 100%. Le contour *bat-17* est récupéré avec un score de 89% : ce score élevé indique que les deux modèles sont semblables comme on peut le constater visuellement. Le

contour suivant reconnu est le contour *bat-01* avec un taux de similarité de 40%. Visuellement, les deux contours sont très semblables mais *bat-01* possède quelques détails sur la partie supérieure des ailes qui n'existent pas sur *bat-16* : ce qui explique ce faible score. Les contours *bat-05* et *bat-04* présentent des scores très faibles (3%), contrairement à *bat-01*, ces contours diffèrent par les oreilles devenues ici très apparentes : ce qui crée une forte dissimilarité avec la requête.

Le tableau III récapitule les résultats qui nous permettent d'effectuer une comparaison de la méthode adaptée de CSS et de la méthode de MSGPR. Dans cette étude, nous avons utilisé trois familles de contours : *bat*, *bell* et *children* extraits de la base de données CE-Shape-1 de MPEG-7. Chaque famille, possède vingt contours d'un même objet représentant différents angles de vue ou déformations dues aux mouvements. Pour une famille donnée, nous utilisons chacun de ses vingt contours comme requête et puis enregistrons les scores minimum et maximum de similarité ainsi que le nombre de contours retenus à la première étape de présélection. La valeur moyenne de chacun de ces trois paramètres est alors consignée dans ce tableau.

Pour chaque méthode, la valeur moyenne des scores maximum de similarité est 100% indiquant donc que le contour de requête est toujours identifié. Ce paramètre indique l'aptitude d'une méthode à reconnaître et à identifier un modèle quand celui-ci existe dans la base de données.

Par ailleurs la valeur moyenne des scores minimum de similarité ainsi que la valeur moyenne du nombre de modèles retenus indiquent la capacité de la méthode à identifier un modèle ayant subi des déformations : plus ces paramètres seront élevés et meilleure sera la méthode. Comme on peut l'observer sur le tableau VII.3, les deux paramètres semblent se comporter de façon antagoniste l'un de l'autre. En effet, tandis que la méthode basée sur la carte CPI fournit le nombre moyen global le plus élevé (14.63) et la moyenne globale des scores minimum la plus basse (19.77), la méthode basée CSS fournit un nombre moyen global faible (6.60) et une moyenne globale des scores minimum plus élevée (47.20). Ceci s'explique par le fait que plus la population retenue est grande et plus il y aura de diversité parmi les individus. Prise sous cet angle, notre approche se montre plus "prudente" en réduisant le nombre d'individus à retenir afin de réduire les risques d'erreurs.

Méthode	Contour	Moyenne des scores maximum	Moyenne des scores minimum	Moyenne du nombre de contours récupérés
CPI	<i>Bats</i>	100.00	36.66	3.90
	<i>Bell</i>	100.00	10.85	20.00
	<i>Children</i>	100.00	11.55	20.00
	Moyenne globale	100.00	19.77	14.63
CSS	<i>Bats</i>	100.00	50.93	4.60
	<i>Bell</i>	100.00	42.52	4.50
	<i>Children</i>	100.00	48.17	10.70
	Moyenne globale	100.00	47.20	6.60

Tableau VII.3 : Etude comparative entre notre méthode et la méthode basée CSS

VII.3.8.2.4 - Observations sur les résultats

Face à ces résultats, nous pouvons conclure que les attributs dérivés de la carte CPI s'affirment avec un bon niveau de discrimination face à ceux de la CSS et révèlent également, des points prometteurs. Bien qu'elle présente quelques insuffisances en présence des déformations élastiques et des occultations, l'approche basée sur la carte CPI possède de sérieux atouts face aux déformations rigides et ce même en présence du bruit.

VII.3.8.3 - Application à la lecture automatique de plaques minéralogiques

Il s'agit de réaliser une lecture automatique des plaques minéralogiques à partir de leurs images. Pour étudier cette problématique, nous appliquons notre approche à la reconnaissance automatique de caractères inscrits sur la plaque minéralogique. La figure 34 montre deux images de plaques réalisées avec des polices de caractères différentes : la différence apparaît plus nettement pour le chiffre '3' sur les deux plaques.



Fig.VII.34 : Exemples d'images de plaques minéralogiques.

VII.3.8.3.1 - Procédure de reconnaissance de caractères

Trois étapes vont structurer notre approche de lecture automatique des plaques minéralogiques. Ces étapes sont récapitulées sur la figure 35 :

- (I) **détection de contours** afin d'obtenir les contours délimitant chaque caractère dans l'image (Fig.VII.30.a et Fig.VII.30.b). On doit noter que cette étape s'avère très importante dans notre processus car de sa précision dépendra l'efficacité de la reconnaissance de caractères. En effet, comme on l'a vu dans des études précédentes, notre méthode est sensible aux transformations élastiques et/ou à la présence d'un bruit fort.
- (II) **extraction du contour caractéristique** de chaque caractère (Fig.VII.30.c). Dans cette étape, on considère seulement le contour externe car dans l'état actuel de notre étude, seuls ces contours sont pris en compte. Comme pour l'étape (I), il est nécessaire d'accorder beaucoup de soin à l'extraction des caractères de sorte qu'ils soient continus et fermés, sans auto-intersection.
- (III) **reconnaissance de caractères**. A cette dernière étape, nous appliquons notre méthode d'analyse pour extraire les attributs de la carte CPI et les intégrer dans le processus d'identification mesurant les taux de similarité entre chaque caractère extrait (et utilisé comme requête) et les modèles présents dans la base de données, en suivant la démarche décrite dans les sections *Choix des attributs* et *Mesure de similarité*.

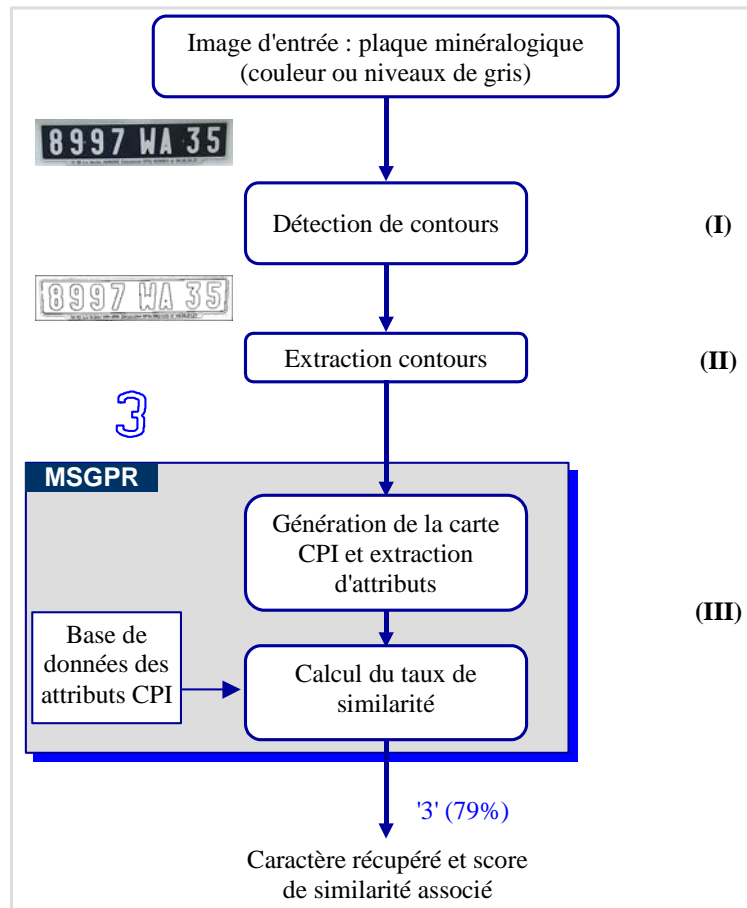


Fig.VII.35 : Vue d'ensemble de notre système de lecture automatique de plaques minéralogiques.

VII.3.8.3.2 - Résultats expérimentaux

Les figures VII.36.a et VII.36.b représentent les images de sortie du détecteur de contours correspondant aux images d'entrée de la figure VII.28. La figure VII.36.c présente l'ensemble des caractères extraits des figures VII.36.a et VII.36.b. Sur la figure VII.36.d nous présentons quelques échantillons de caractères de la base de données : cette base consiste en un ensemble de caractères extraits du jeu de caractères "bold.chr" de Borland®.

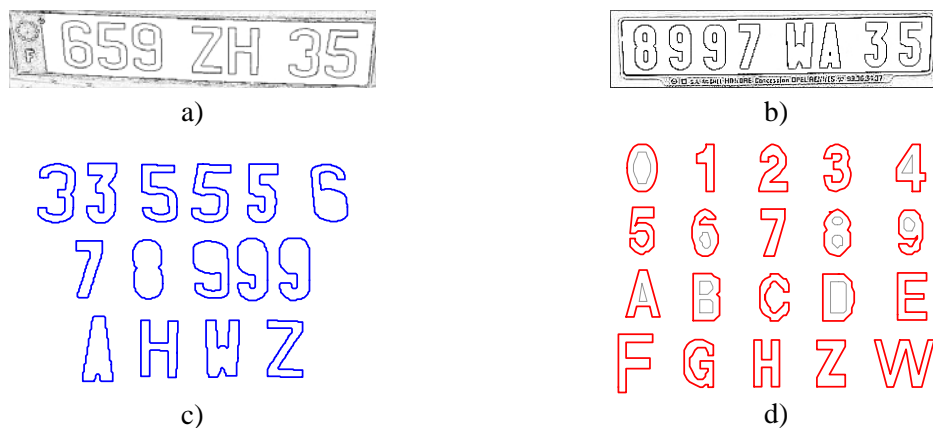


Fig.VII.36 : a) et b) Contours détectés - c) contours extraits de a) et de b) - d) modèles de contours dans la base de données.

On notera que dans cette étude, la base de données se compose d'une seule police de caractères : tandis que les caractères requête à reconnaître sont de deux polices différentes. Pour améliorer les résultats d'identification des caractères, une solution serait l'intégration dans la base de données, de toutes les polices possibles utilisées pour l'immatriculation des véhicules.

La figure VII.37 montre quelques résultats obtenus à partir des images d'entrée de la figure VII.28.

Sur ces figures, nous présentons quelques résultats de reconnaissance de caractères. Sur chaque figure, le contour en haut et à gauche représente le contour requête. Les contours suivants dans le sens d'un balayage lexicographique, représentent les huit contours de la base de données qui donnent les scores de similarité les plus élevés : en commençant par les scores les plus forts.

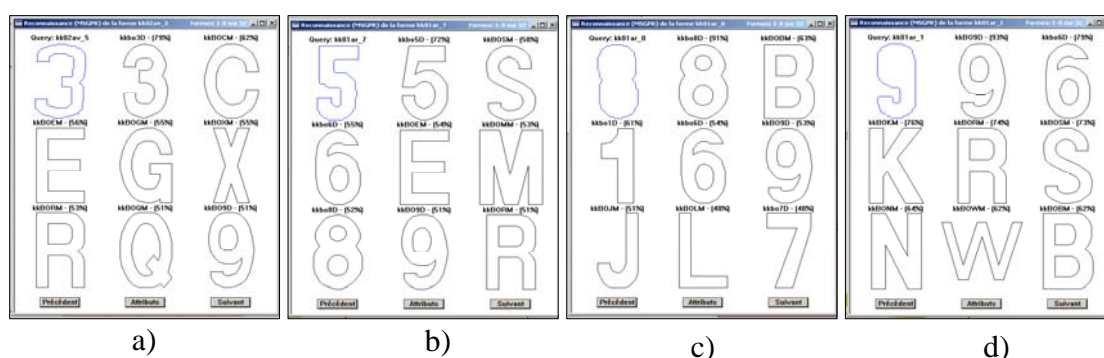


Fig.VII.37 : Exemples de résultats de reconnaissance de caractères

Comme le montrent ces résultats, l'identification des différents caractères est relativement efficace : pour chaque requête, le caractère identifié (la seconde forme de chaque des figures VII.37.a-d) est bien le caractère recherché. Ainsi, pour la requête '3', nous identifions le chiffre '3' avec des scores de similarité de 79%. Le Tableau VII.4 récapitule les trois scores de similarité les plus élevés pour les contours présentés sur la figure VII.37. Pour le chiffre '9' comme requête, nous avons récupéré le chiffre '9' avec un score de similarité de 96% suivi du chiffre '6' avec un score de similarité de 79%. Notons que le chiffre '6' de cette police n'est autre que le chiffre '9' qui a subi une rotation de 180° d'où le score de similarité élevé pour '6'. On pourra noter la ressemblance morphologique entre le chiffre '5' et la lettre 'S' ou entre le chiffre '8' et la lettre 'B'. Ceci se traduit également par des apparitions en seconde position mais, par contre, les spécificités propres à chaque caractère conduisent à des écarts des scores de similarité suffisamment conséquents pour éviter des méprises.

Requête	Lettre reconnue (Score de similarité)		
'3'	'3' (79%)	'C' (62%)	'E' (56%)
'5'	'5' (72%)	'S' (58%)	'6' (55%)
'8'	'8' (91%)	'B' (63%)	'1' (61%)
'9'	'9' (96%)	'6' (79%)	'K' (76%)

Tableau VII.4 : Echantillon de caractères reconnus avec les taux de reconnaissance associés.

VII.3.9 - Conclusion

Nous avons présenté une méthode originale pour la description des contours dans le contexte de la reconnaissance des formes planaires.

Cette approche exploite une analyse multiéchelle du contour en s'appuyant sur un lissage par réduction progressive de la largeur de bande du filtre. La détection des points d'intersection entre le contour original et les contours filtrés permet de générer la carte des points d'intersection (CPI) indiquant leur localisation, dans un plan espace-échelle, et caractérisant la forme.

Les résultats expérimentaux obtenus sur divers contours ont montré que la carte CPI est invariante à la translation, à la rotation et au changement d'échelle pour une large gamme d'échelles. Ainsi, nous avons extrait quelques paramètres intéressants dérivés de la carte CPI pour servir d'attributs caractérisant l'objet planaire concerné.

L'application de ces attributs à la récupération d'une forme dans une base de données et une étude comparative avec la méthode CSS ont donné des résultats encourageants. Ils ont montré l'efficacité et les sérieux atouts de cette approche dans la description et la reconnaissance des formes planaires sans occultation.

Une étude expérimentale appliquant les attributs dérivés de la carte CPI à la lecture automatique de plaques minéralogiques donne des résultats particulièrement intéressants.

La faiblesse essentielle de cette approche est son incapacité à reconnaître les formes partiellement occultées. Cette incapacité est, principalement, due à la structure intrinsèque de la carte CPI qui permet de capturer les informations globales mais pas suffisamment locales. Pour pallier ce problème, nous avons défini une nouvelle approche qui, cette fois-ci, permet de capturer de l'information multiéchelle localisée dans un voisinage limité. Grâce à cette aptitude, la nouvelle approche se présente comme un outil prometteur vis-à-vis de la reconnaissance des formes dont celles qui sont partiellement occultées. Une présentation succincte de cette dernière approche est donnée dans les perspectives de la section suivante : nous pensons que la combinaison judicieuse de ces deux approches permettra d'améliorer grandement les performances de la MSGPR.

VIII - Perspectives

Mes perspectives en matière de recherche se déclinent d'une part sur la description et la reconnaissance des formes et d'autre part sur la fusion d'images de télédétection. Grâce aux deux aspects de l'analyse d'images dans lesquels nous avons acquis de l'expertise, l'objectif est de poursuivre les études pour faire converger les deux aspects vers des solutions plus efficaces dans l'interprétation d'images.

VIII.1 - En analyse multiéchelle

VIII.1.1 - Méthode MSGPR

L'objectif principal dans ce domaine est d'approfondir cette nouvelle méthode d'analyse multiéchelle pour la reconnaissance des formes. La méthode n'étant qu'à ses débuts, il est clair que beaucoup d'études restent encore à mener afin de se fixer sur ses performances et pour déterminer ses limites réelles :

- en approfondissant la détermination des coefficients A_x et A_y dans l'ajustement d'échelle du contour lissé : de cette opération dépend la stabilité de la carte CPI,
- en approfondissant l'étude des structures caractérisant les portions convexes/concaves (par exemple, dans l'analyse des défauts des pièces usinées),
- en développant une version dans laquelle le contour sera représenté, non plus dans l'espace cartésien, mais dans l'espace polaire.

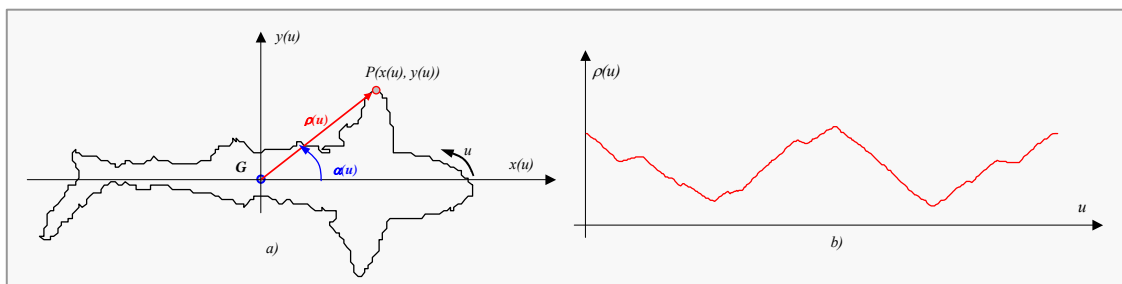


Fig.VIII.38 : Contour dans l'espace polaire et le signal $\rho(u)$ correspondant

Le contour sera alors représenté par l'angle $\alpha(u)$ que fait le point $P(x(u), y(u))$ avec

l'axe des abscisses et le "rayon" $\rho(u)$ défini par la distance entre le centre de gravité \mathbf{G} et le point \mathbf{P} comme on peut le voir sur la figure VIII.38.

VIII.1.2 - Approche basée sur l'angle tangentiel

VIII.1.2.1 - Introduction

Comme nous l'avons indiqué plus haut, cette nouvelle approche permet d'extraire des informations plus localisées que ne le sont celles provenant de la carte CPI de MSGPR. De la même façon que pour MSGPR, le contour d'entrée est défini par une liste de N points représentés par leurs coordonnées dans le plan. Pour la normalisation, nous avons gardé le même nombre de points ($N=360$). Les points du contour sont triés et ordonnés dans le sens trigonométrique direct et le contour est paramétré à l'aide de l'abscisse curviligne u exactement comme dans le cas de MSGPR. Etant donné un point initial P_0 , les points du contour sont indexés de P_0 à P_{N-1} . Soit alors, \vec{V}_n le vecteur défini par deux points consécutifs P_n et P_{n+1} orienté dans le sens P_n vers P_{n+1} . Dans cette configuration, on peut donner les définitions suivantes :

Définition 1 : l'angle θ_n mesuré dans le sens trigonométrique direct au point P_n entre l'axe des abscisses et le vecteur \vec{V}_n est appelé angle tangentiel (ou turning angle).

Définition 2 : l'angle φ_n mesuré dans le sens trigonométrique direct entre deux vecteurs consécutifs \vec{V}_{n-1} et \vec{V}_n correspondant à l'écart entre deux angles tangentiels consécutifs est l'angle tangentiel différentiel (differential-Turning Angle : d-TA)

$$\varphi_n = \theta_n - \theta_{n-1} \quad (23)$$

Ce paramètre permet d'obtenir une estimation de la courbure d'un contour. En fonction de son signe, il permet également de déterminer si la partie concernée du contour est une zone concave ou convexe.

La figure VIII.39 illustre une configuration indiquant les différents paramètres. Comme on peut le noter sur cette figure, les angles mesurés sur des points d'une portion concave fournissent des valeurs négatives (voir φ_k sur la figure) alors que ceux obtenus sur des points de portions convexes donnent des valeurs positives (voir φ_n sur la figure).

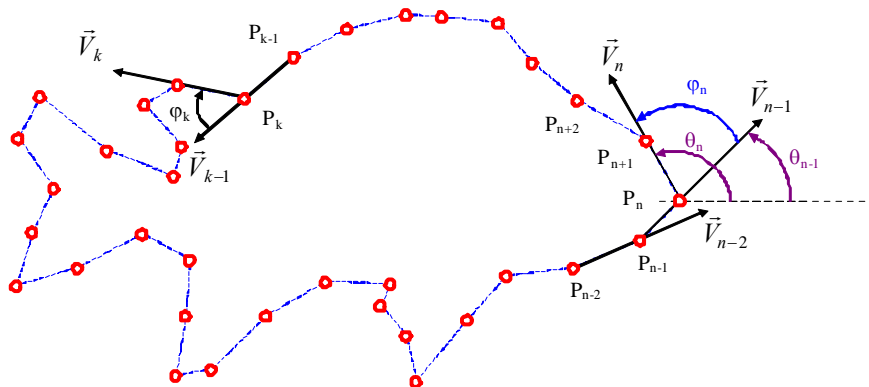


Fig.VIII.39 : Exemple de contour illustrant l'angle tangentiel

VIII.1.2.2 - Motivations

Dans [Feldman05], on cite F. Attneave qui, déjà en 1954, s'était intéressé au problème et avait suggéré que l'information visuelle le long d'un contour était concentrée dans des régions à forte courbure mais pas uniformément distribuée sur le contour. Pour illustrer ses propos, il a montré qu'un chat dormant (voir la figure VIII.40 ci-contre) pouvait être décrit à l'aide de 38 points seulement. Ces points de plus forte courbure sont joints par de simples segments de droite. Partant de cette idée, J. Feldman et al. ont observé que le signe de la courbure apportait aussi de l'information utile [Feldman05]. Ainsi, ils ont pu observer que les courbures négatives (les concavités) étaient beaucoup plus riches en information que les courbures positives (les convexités). De cette observation, il est raisonnable de penser que la fonction d-TA est un outil adapté pour extraire de l'information utile à l'analyse d'un contour. La figure VIII.41 montre un contour et l'angle tangentiel exprimé en fonction de l'abscisse curviligne u . Les lettres (a), (b),..., (k) sur le contour indiquent les parties correspondantes (A), (B),..., (K) sur la fonction d-TA(u). On observe ainsi que les zones à courbure constante (c) et (i) produisent une fonction d-TA constante (C) et (I). (c) qui est une convexité produit une fonction d-TA positive alors que (i) qui est une concavité produit une d-TA négative. De même, on peut observer que les zones à forte convexité (a) ou à forte concavité (b) produisent une forte valeur positive (A) et respectivement négative (B) de la fonction d-TA. Les segments de droites (f) conduisent, eux, à un angle tangentiel différentiel nul (F). De cette observation, nous définissons les points essentiels qui peuvent être regroupés en trois catégories selon le signe de leur courbure.



Définition 3 : Les points essentiels sont les points correspondant aux valeurs extrêmes ou les passages par zéro de la fonction d-TA. On observe ainsi, les points suivants :

- **points α** localisés aux minima de la fonction angle tangentiel différentiel d-TA,
- **points β** localisés aux maxima de la fonction angle tangentiel différentiel d-TA,
- **points γ** localisés aux passages par zéro de la fonction angle tangentiel différentiel d-TA.

A ce niveau, on peut rappeler que la fonction d-TA n'est autre qu'une autre façon d'estimer la courbure et que de ce fait les points γ sont une représentation des points de la carte des passages par zéro de la courbure exploitée dans les approches CSS [Mokhtarian92].

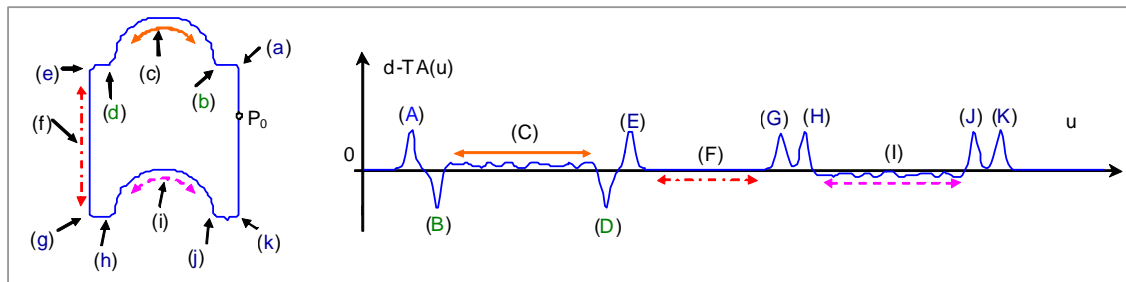


Fig.VIII.41 : Exemple de contour et la fonction angle tangentiel.

VIII.1.2.3 - Analyse multiéchelle

Afin de capturer le maximum d'information et bénéficier de la richesse d'une analyse multiéchelle, un lissage progressif est appliqué au contour. Ce lissage est réalisé à l'aide d'un filtrage passe-bas appliqué aux coordonnées paramétriques $x(u)$ et $y(u)$: on obtient ainsi les coordonnées paramétriques $X(u,\sigma)$ et $Y(u,\sigma)$ pour le point après lissage. Ici, le filtrage se fonde sur l'itération d'un noyau gaussien d'écart-type $\sigma_0 = \sqrt{2/\pi} = 0.7979$ et de longueur 3 ($g = [0.25, 0.50, 0.25]$). En réitérant n fois l'action de filtrage, il revient à appliquer un filtrage avec un noyau gaussien d'écart-type

$$\sigma_n = \sigma_0 \sqrt{n} \quad (24)$$

En représentant la fonction d-TA des contours filtrés dans un plan espace-échelle, on obtient le scalogramme de la figure VIII.42.a où l'ordonnée représentée par le nombre d'itérations indique le paramètre d'échelle σ conformément à l'équation (24). L'abscisse, quant à elle, représente le paramètre spatial u échantillonné sur 360 points.

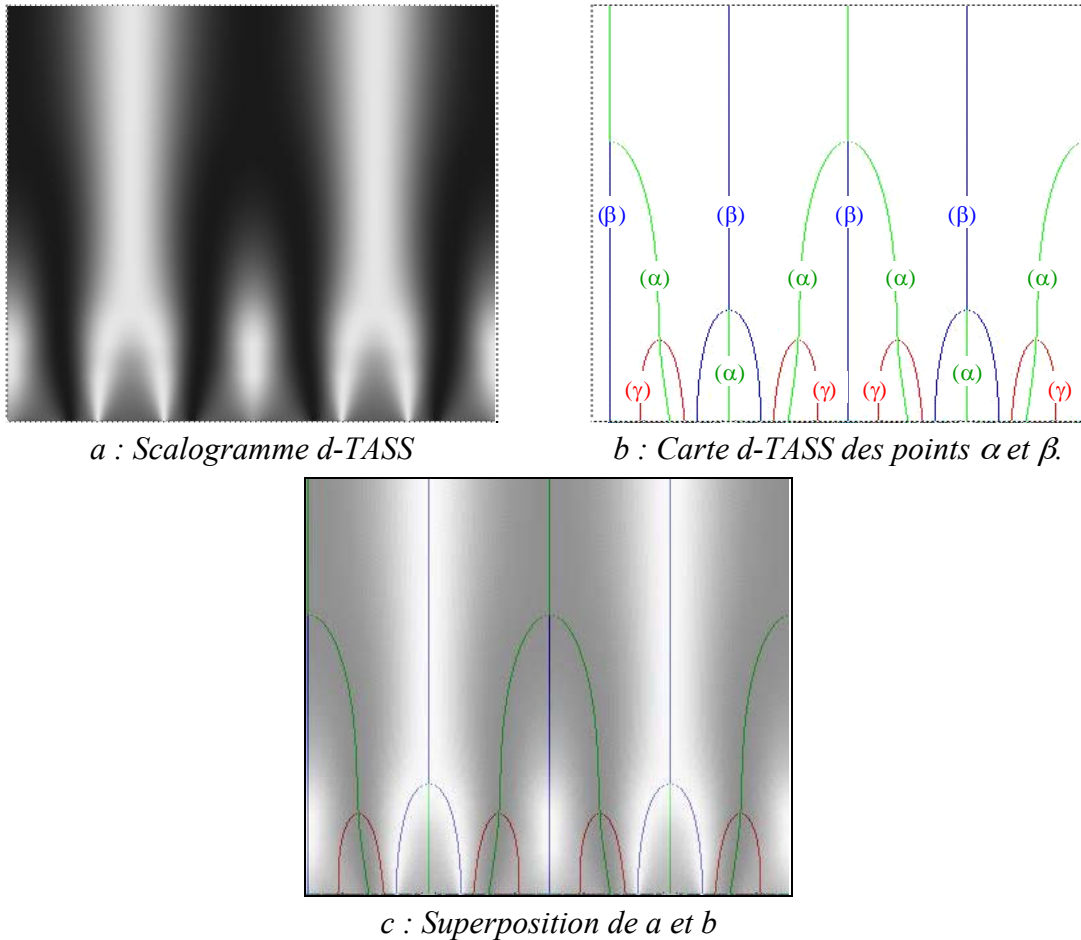


Fig.VIII.42: Scalogramme d-TASS et la carte de points essentiels

La valeur du scalogramme au point (u,σ) indique alors l'angle tangentiel au point $P(X(u,\sigma), Y(u,\sigma))$ obtenu après lissage du contour original par un filtre gaussien de d'écart-type σ_n . Ce scalogramme est dénommé fonction d-TASS pour "*differential-Turning Angle Scale Space function*". La figure VIII.42.b représente la carte des points essentiels α , β et γ extraits de cette fonction d-TASS : c'est la carte d-TASS. La

superposition des deux entités sur la figure VIII.42.c montre clairement la localisation des différents points essentiels dans le scalogramme d-TASS :

- les points α sont localisés aux minima de la fonction d-TA (zones sombres),
- les points β aux maxima de la fonction d-TA (zones claires) et
- les points γ aux passages par zéro de la fonction d-TA.

Il faut noter que les points α ne correspondent pas uniquement aux points à minima négatifs de d-TA. Tous les minima sont pris en compte : ce qui explique que les points α et β persistent encore alors que les points γ se sont estompés à une échelle Σ_0 donnée. Ces points α et β vont continuer tant qu'il y aura des extrema : c'est-à-dire tant que le contour lissé n'est pas égal à un cercle.

Remarque :

En fonction des choix, la condition d'arrêt peut être l'estompement total des points essentiels γ : dans ce cas-ci, on ne conserve donc pas les minima positifs. La figure VIII.43 décrit un exemple de la procédure de génération de la carte d-TASS.

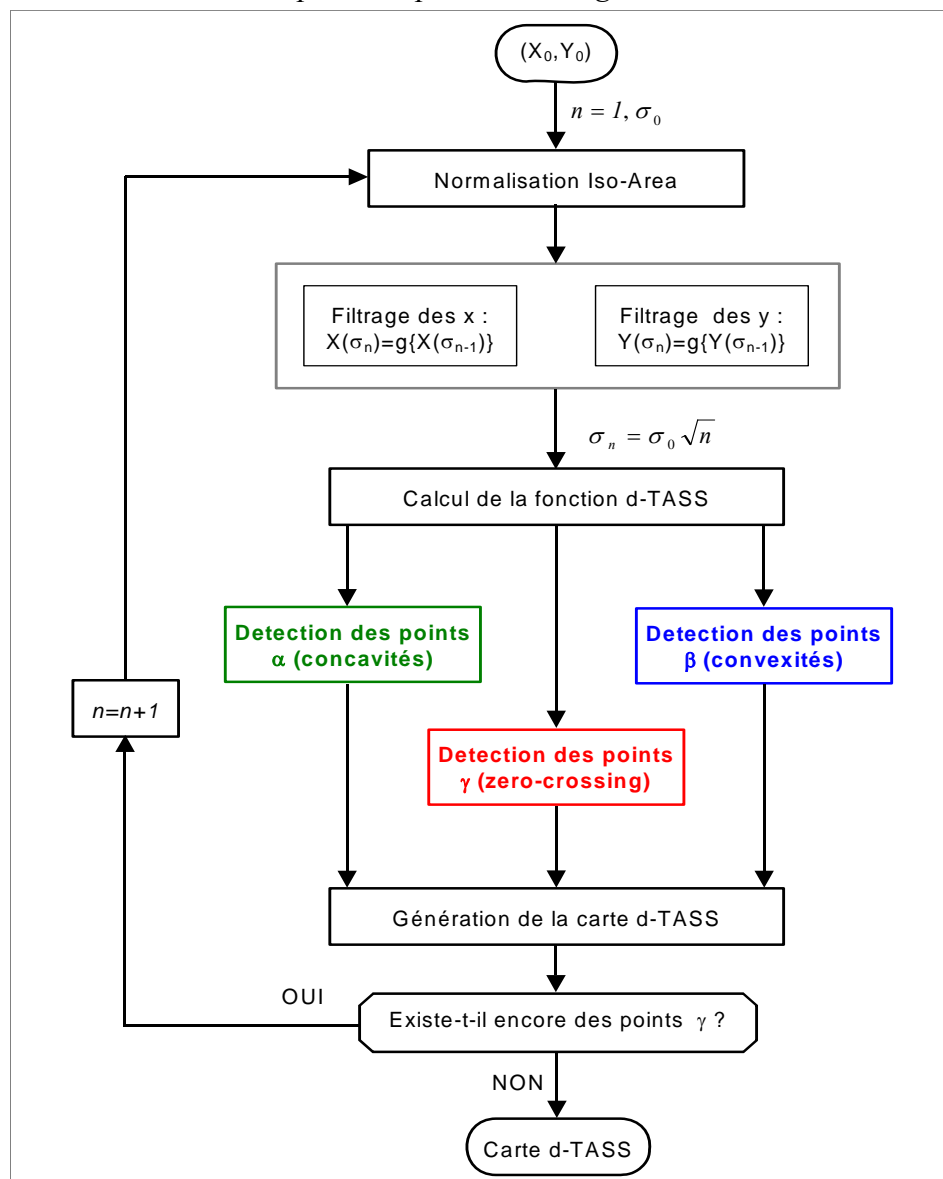


Fig.VIII.43: Procédure de génération de la carte d-TASS

Des études préliminaires nous ont permis de montrer que le scalogramme d-TASS est caractéristique du contour concerné [Kpalma07a, Kpalma07b]. Il présente une série de propriétés très intéressantes qui constituent les propriétés essentielles dans le choix d'un descripteur des formes :

- **représentativité des formes (voir section Description de formes par analyse multiéchelle de l'angle tangentiel)** : la fonction (ou la carte) d-TASS est caractéristique du contour concerné, car pour deux contours différents, nous obtenons deux fonctions (ou deux cartes) d-TASS différentes.
- **résistance à la rotation** : la fonction (ou la carte) d-TASS est topologiquement insensible à la rotation (excepté une translation circulaire qui peut être facilement compensée).
- **résistance au changement d'échelle** : la fonction (ou la carte) d-TASS reste pratiquement identique pour une large gamme d'échelle,
- **résistance à l'occultation** : la fonction (ou la carte) d-TASS permet d'extraire des structures locales représentant des portions du contours : pour cela, elle est capable de reconnaître des objets partiellement occultés,
- **résistance à la transvection** : la fonction (ou la carte) d-TASS est topologiquement insensible à la transvection qui est une expression de la projection en perspective liée au changement de l'angle de vue.

Avec ces informations à l'esprit, la tâche suivante et non la moindre est d'extraire des scalogrammes (la fonction ou la carte d-TASS) les mesures adéquates pour donner un descripteur efficace.

VIII.1.2.4 - Définition d'un descripteur : quelques pistes

Avant de définir un descripteur, il est nécessaire d'analyser les scalogrammes obtenus : les figures VIII.44 et VIII.47 montrent des exemples de blocs réalisés, d'une part, dans la fonction d-TASS et, d'autre part, dans la carte d-TASS. On observe que, pour un bloc donné, la moyenne du contenu est caractéristique de la forme. La figure VIII.48 illustre les pics des points essentiels. On observe que leur nombre et leur localisation dans la carte sont caractéristiques du contour.

De ces observations, il convient de définir des mesures qui soient caractéristiques et discriminantes pour décrire une forme. Dans l'état actuel de l'étude, bien que certaines idées soient plus ou moins bien avancées, nous ne pouvons que donner quelques pistes dans la définition d'un descripteur à partir des scalogrammes d-TASS. Il s'agit 1) de sous-échantillonner la carte d-TASS en blocs et extraire une mesure par bloc, 2) réaliser cette opération sur la fonction d-TASS, 3) utiliser les coordonnées des pics de la carte d-TASS ou 4) utiliser les distances entre points essentiels à une série d'échelles donnée.

VIII.1.2.4.1 - Solution N° 1 : carte d-TASS sous-échantillonnée

Comme le montre la figure VIII.44, on pourrait utiliser la **carte d-TASS sous-échantillonnée** : découper la carte d-TASS en L lignes et K colonnes puis calculer une fonction dans chaque bloc. On obtient ainsi une matrice attribut m_A à $L \times K$ éléments. Pour la définition du descripteur à base de la carte d-TASS, on utilise une carte intermédiaire qui est une combinaison linéaire des points essentiels. Le descripteur utilisé est alors défini comme suit :

$$I = w_\gamma \cdot I_\gamma + w_\alpha \cdot I_\alpha + w_\beta \cdot I_\beta \quad (25)$$

où le coefficient w_a indique le poids appliqué aux points essentiels $a=\alpha, \beta$ ou γ . I_a correspondant à l'ensemble des points essentiels de type a . Le descripteur mA utilisé est défini comme suit :

$$mA(l,k) = f(I(l,k)) \quad (26)$$

avec $I(l,k)$ le bloc de la ligne l et de la colonne k dans la découpe de la carte I .

Les fonctions f utilisées ici pour calculer l'élément $mA(l,k)$ du descripteur sont les moments d'inertie m_{00} , m_{01} et m_{10} du bloc correspondant.

Comme le témoignent les résultats ci-dessous, obtenus dans une étude préliminaire fondée sur cet aspect, les attributs de cette approche offrent d'assez bonnes performances.

Pour évaluer cette approche, nous avons constitué une base de contours comme suit : nous avons sélectionné cinquante contours de la base CVSSP auxquels nous avons appliqué une transformation affine composée d'une rotation d'angle θ , une homothétie h et une transvection λ dont les valeurs sont indiquées dans le tableau VIII.5. Dans ce tableau, nous avons fixé arbitrairement, trois valeurs par transformation élémentaire.

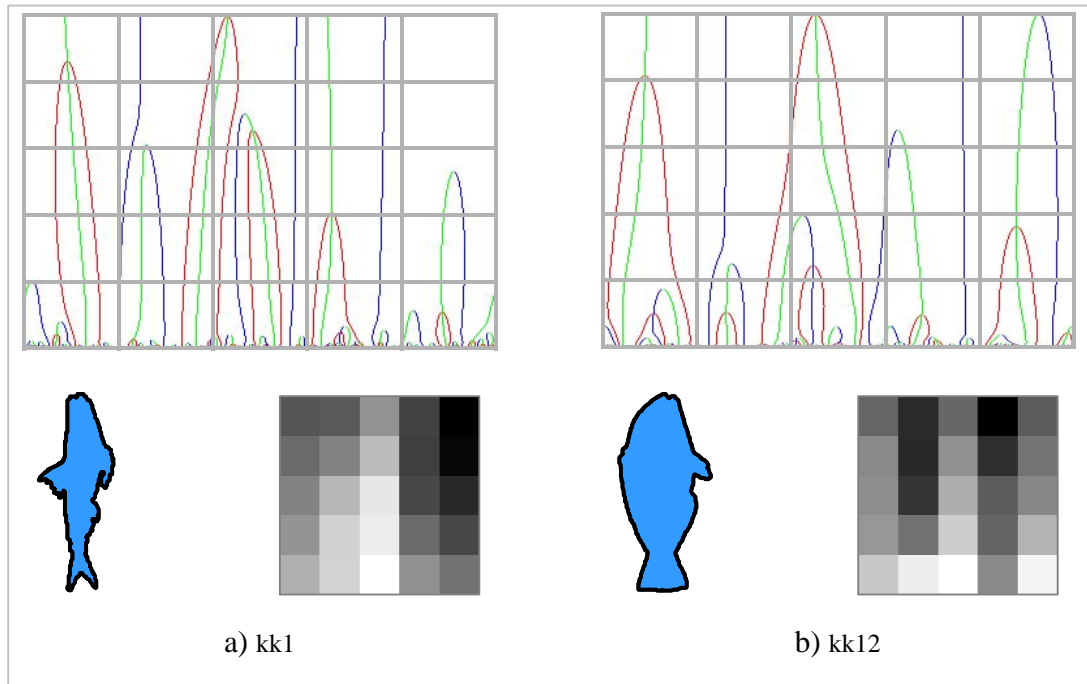


Fig.VIII.44 : Attributs issus d'un sous-échantillonnage de la carte d-TASS

Ce qui fait que pour un contour original donné, nous avons une famille comportant un total de $3^3+1=28$ individus. Pour une requête donnée, nous avons donc 28 contours pertinents (ou vrais positifs) dans la base. Ceci donne un total de 1400 individus dans la base.

		Indices n des paramètres		
		0	1	2
Paramètres	θ_n	45°	60°	90°
	h_n	0.25	0.75	5
	λ_n	0.3	0.6	1.5

Tableau VIII.5 : Ensemble des paramètres de transformation

La figure VIII.45 montre l'ensemble des individus de la famille kk1 après les avoir remis à l'échelle. Pour l'évaluation, nous présentons chacun des contours de la base comme requête et nous mesurons le rappel et la précision ainsi que les paramètres BEP (Bull's eye performance) et AUC (Area under R-P curve).

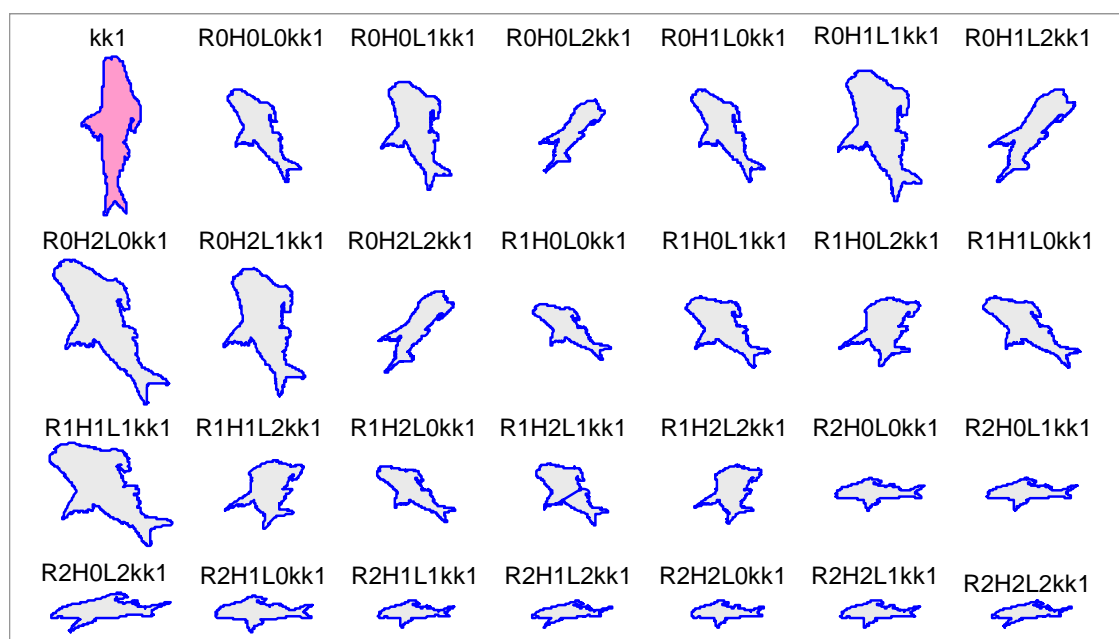


Fig.VIII.45 : Exemple de contour avec les transformations affines correspondante

Sur la figure VIII.46, on peut voir l'allure des courbes de Rappel-Précision (R-P) pour une série de combinaisons des points essentiels et pour des tailles différentes du descripteur. On observe ainsi que certaines combinaisons donnent des performances très intéressantes : celles dont les courbes R-P épousent le bord supérieur et le bord droit du graphique. Ces observations sont confirmées par les valeurs élevées des paramètres BEP et AUC du tableau VIII.6. Plus ces paramètres sont proches de l'unité et plus le système de recherche est considéré comme performant. Dans cette étude, les blocs ont été réalisés avec un recouvrement de 25%.

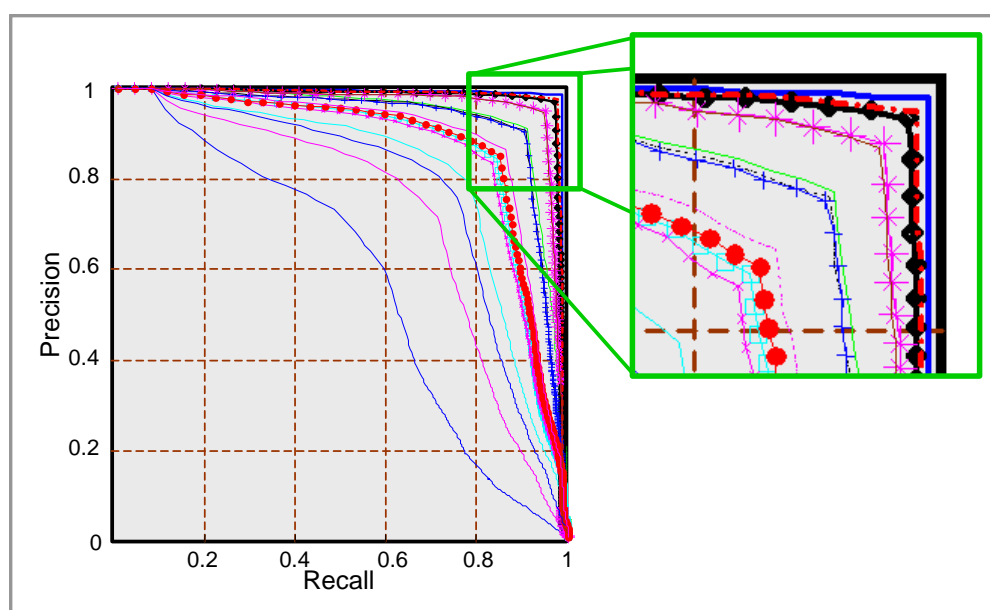


Fig.VIII.46 : Graphes de Rappel-Précision pour 16 combinaisons des points essentiels

Poids des points essentiels	Mesure de performance	Taille de la matrice attribut mA			
$[w_\gamma \ w_\alpha \ w_\beta]$		4x4	8x8	16x16	32x32
1 [1 0 0]	BEP AUC	0.986 0.964	0.986 0.963	0.986 0.963	0.986 0.963
2 [0 1 0]	BEP AUC	0.887 0.811	0.926 0.861	0.969 0.921	0.981 0.946
3 [0 0 1]	BEP AUC	0.758 0.647	0.669 0.552	0.697 0.580	0.720 0.604
4 [0 1 1]	BEP AUC	0.845 0.753	0.793 0.690	0.814 0.717	0.840 0.748
5 [1 1 1]	BEP AUC	0.955 0.916	0.930 0.867	0.917 0.857	0.916 0.854
6 [0.5 0.25 0.25]	BEP AUC	0.982 0.953	0.961 0.922	0.956 0.911	0.959 0.909
7 [0.75 0.125 0.125]	BEP AUC	0.986 0.964	0.983 0.959	0.985 0.957	0.982 0.951
8 [0.2 0.4 0.4]	BEP AUC	0.919 0.860	0.879 0.799	0.876 0.798	0.885 0.813
9 [1 -1 1]	BEP AUC	0.933 0.893	0.899 0.833	0.913 0.846	0.941 0.879
10 [1 -0.5 0.5]	BEP AUC	0.972 0.942	0.954 0.910	0.956 0.909	0.966 0.919
11 [1 -0.25 0.25]	BEP AUC	0.982 0.961	0.975 0.948	0.976 0.944	0.979 0.943
12 [1 1 -1]	BEP AUC	0.938 0.897	0.915 0.857	0.925 0.862	0.935 0.876
13 [1 0.125 -0.125]	BEP AUC	0.986 0.964	0.986 0.962	0.985 0.960	0.982 0.951
14 [1 0.25 -0.25]	BEP AUC	0.983 0.960	0.977 0.949	0.975 0.943	0.978 0.940
15 [0.75 0.5 1]	BEP AUC	0.911 0.857	0.856 0.773	0.850 0.768	0.860 0.780
16 [0.75 1 0.5]	BEP AUC	0.948 0.900	0.935 0.877	0.930 0.875	0.938 0.879
Moyenne	BEP AUC	0.936 0.890	0.914 0.858	0.919 0.863	0.928 0.872

Tableau VIII.6 : Exemple de performances pour 16 combinaisons des points essentiels

Le tableau VIII.6 et la figure VIII.47 indiquent, tous les deux, que les performances varient en fonction de la pondération affectée aux points essentiels dans leur

combinaison. On notera, néanmoins, que même tout seuls, les points γ donnent de très bonnes performances.

On observe, également, une variabilité des performances en fonction de la taille de la matrice d'attributs. Toutefois, comme le montre la ligne "Moyenne" du tableau, il faut noter que la taille de 4×4 semble apporter des performances satisfaisantes sans avoir besoin de recourir à des tailles plus importantes.

VIII.1.2.4.2 - Solution N° 2 : fonction d-TASS sous-échantillonnée

Au lieu d'utiliser la carte d-TASS comme précédemment, on pourrait utiliser la **fonction d-TASS sous-échantillonnée** : découper la carte d-TASS en L lignes et K colonnes puis calculer une fonction dans chaque bloc (voir équation (26)). On obtient ainsi une matrice attribut à $L \times K$ éléments comme dans le cas précédent.

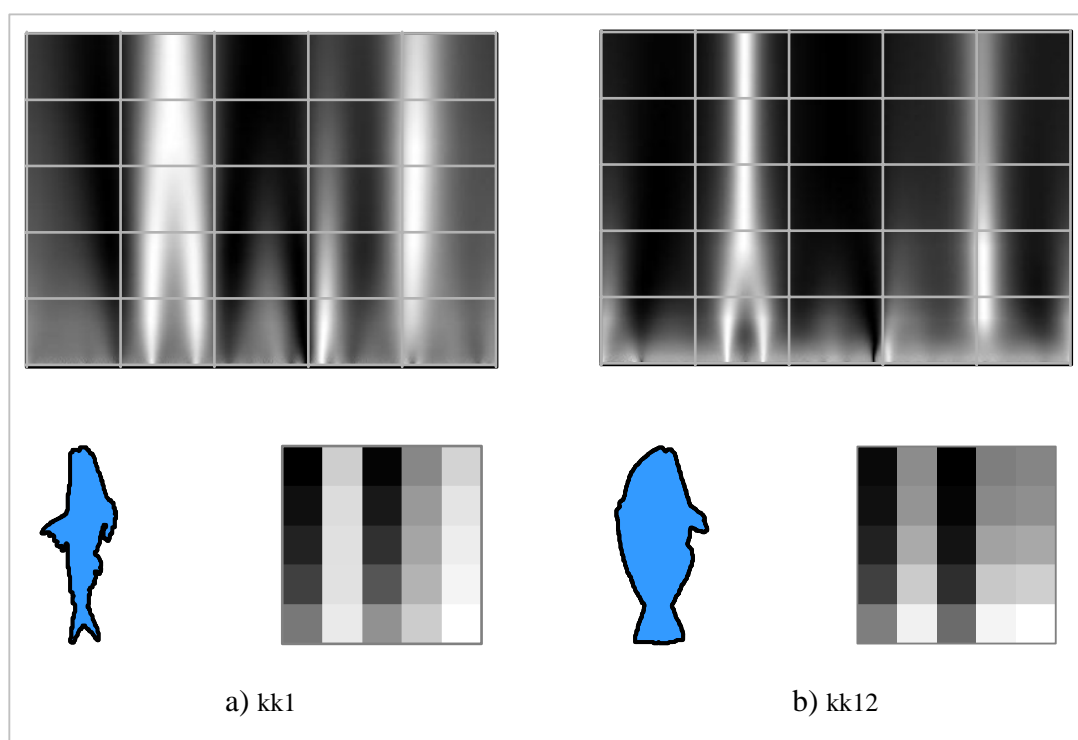


Fig.VIII.47 : Attributs issus d'un sous-échantillonnage de la fonction d-TASS

VIII.1.2.4.3 - Solution N° 3 : pics de la carte d-TASS

Ici, on propose d'utiliser les **pics des points essentiels** : pics des points α et β combinés, ou pics des points γ à la manière de CSS ou tous les pics confondus (voir figure VIII.48). Les coordonnées (u, σ) de chaque pic constitueraient ainsi une base pour la définition du descripteur correspondant.

En exploitant cet aspect conjointement avec la notion de matrice de distance, nous avons obtenu des résultats très intéressants [Kpalma08]. Dans cette approche, la requête est utilisée comme référence. Ainsi, pour chaque modèle de la base, nous apparions les pics de la requête aux pics les plus proches dans la carte de la référence. Lorsque la distance minimale entre deux pics est plus grande que le seuil préfixé, le point homologue correspondant est mis à zéro.

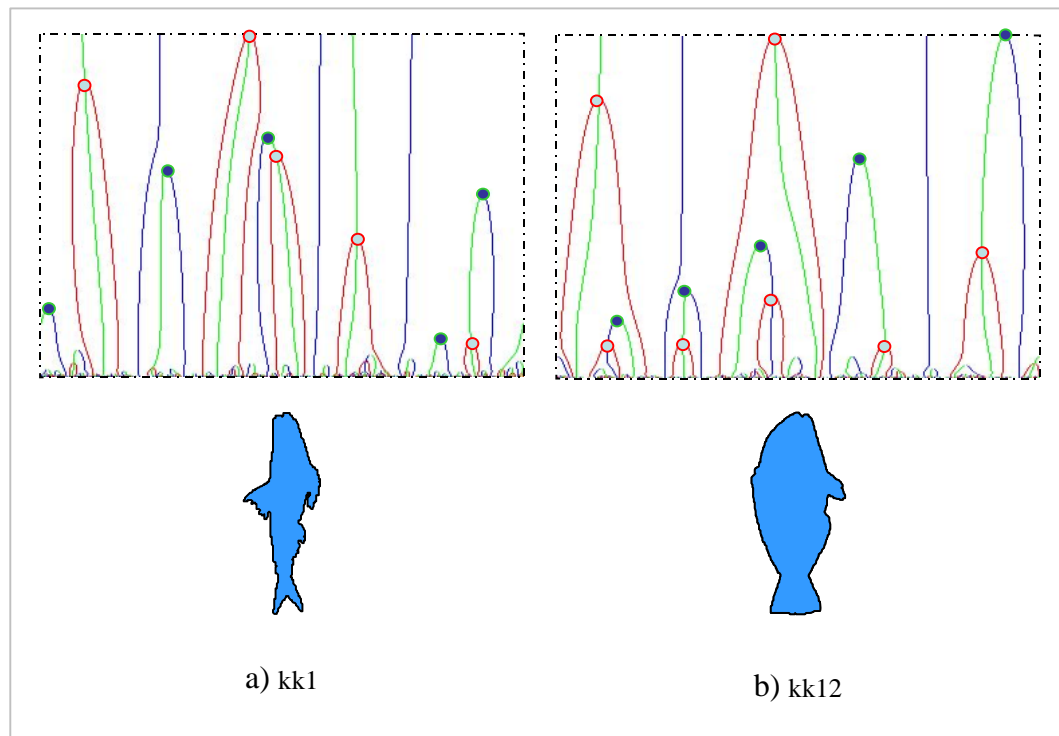


Fig.VIII.48 : Attributs basés sur les pics des points essentiels

VIII.1.2.4.4 - Solution N° 4 : distance entre les points essentiels par échelle

A la manière de la méthode MSGPR, on propose de fixer une série d'échelles et d'utiliser les **écarts entre deux points consécutifs à une échelle donnée** (voir la figure VII.25). En combinant les mesures extraites à différentes échelles, on construit alors un attribut multiéchelle.

VIII.1.2.5 - Conclusion

Il est évident qu'il est trop tôt pour tirer une conclusion sur les performances de cette nouvelle approche. Elle est appelée à évoluer et c'est après une étude approfondie appuyée par des résultats expérimentaux que nous pourrions conclure sur son évaluation. Une étude comparative entre les deux approches (carte CPI, fonction et carte d-TASS) nous permettra alors de dégager les points forts et les points faibles de chacune afin d'orienter le choix de l'utilisateur. Ensuite, on pourra envisager de définir un attribut conjoint par combinaison des deux approches.

Nos recherches évolueront, ensuite, vers la définition de systèmes d'indexation destinés à la recherche d'images par le contenu comme il est défini par les systèmes CBIR. Ceci nous permettra d'enrichir notre plate-forme de démonstration LAMEFor en cours de développement (voir section Etude des performances en reconnaissance des formes).

VIII.2 - En fusion d'images

Nos travaux en analyse pour la fusion d'images satellitaires ont suivi une évolution constante depuis le filtrage de speckle sur les images radar, la segmentation, le recalage d'images multicanaux et multidates jusqu'à la fusion pour la visualisation. L'objectif est

d'approfondir cette étude afin d'aboutir à un schéma de fusion plus complet comme il est suggéré sur la figure VIII.49 suivante.

Dans cette configuration semi-automatique, nous disposons des boucles de retour pour le contrôle de certaines performances. Toutefois, le choix est laissé à l'opérateur interprète d'images qui possède une expertise et des connaissances a priori, d'influer sur la décision finale.

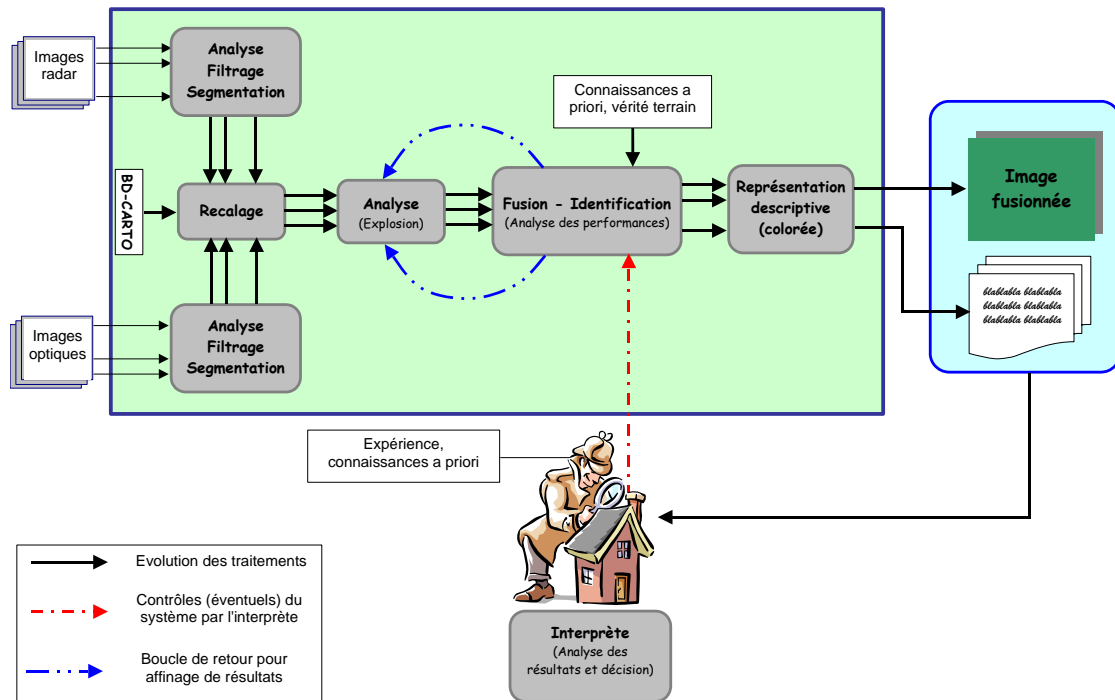


Fig.VIII.49 : Vue synoptique du schéma de fusion consensuelle d'images satellitaires

IX - Conclusion générale

Ce document survole l'ensemble de mes activités de recherche et d'enseignement ainsi que mes diverses responsabilités.

Issu d'une formation dans le domaine de l'électronique et de l'informatique industrielle, mes enseignements et mes recherches restent fortement liés à ces spécialités. Ces activités se situent principalement dans le domaine de l'analyse et de l'interprétation des images : sous les deux aspects de la description des formes et de la fusion d'images de télédétection. Nos travaux dans ce domaine visent au développement de nouvelles méthodes ou à l'adaptation de méthodes existantes pour une meilleure efficacité.

En **description des formes**, les objectifs sont d'étendre nos recherches et de les adapter pour des applications à l'indexation des images par le contenu, ainsi qu'à la détection et au suivi d'objets dans les séquences vidéos. Nous avons développé une méthode originale d'analyse multiéchelle basée sur le filtrage progressif du contour de la forme à analyser. La carte des points d'intersection (CPI) qui en découle est caractéristique de la forme et a permis d'obtenir des résultats intéressants dans le contexte de la description et de la reconnaissance des formes. Une autre étude, réalisée dans le cadre d'une thèse, se fondait sur l'aspect région de l'objet. Inspirée de la transformée de Radon, cette approche exploite une distribution de longueurs de plage par orientation. Une nouvelle étude basée sur la notion de l'angle tangentiel (d-TASS) vient enrichir ces études dans le choix d'un descripteur efficace et compact. La synthèse de ces différentes approches nous permettra de définir un descripteur plus complet prenant en compte les deux entités définissant un objet et qui sont la région et le contour.

Cette solide expérience dans le domaine de l'analyse pour la description des formes, nous conduira vers des solutions hybrides basées aussi bien sur les textures que sur les formes (contours et/ou régions). C'est dans ce cadre que s'inscrit la nouvelle thèse intitulée "Analyse d'images compressées pour leur indexation et pour leur interprétation". Les solutions qui en découleront, trouveront leur place dans les systèmes de "recherche d'images par le contenu", les systèmes de "détection, reconnaissance et identification", les systèmes de surveillance, les systèmes d'authentification...

En **fusion d'images**, l'objectif vise à fournir des moyens techniques pour une assistance à la gestion de l'environnement et permettre un meilleur suivi de l'évolution d'un

écosystème pour une évaluation de l'impact de ses changements. Nous nous sommes intéressés aux problématiques connexes comme le filtrage de speckle, la segmentation et le recalage d'images. Dans ce contexte, nous avons réalisé une fusion d'images pour la visualisation en exploitant la notion d'explosion algorithmique sur une image radar SAR, ceci pour rehausser les forts réflecteurs par une analyse associant la vision locale et la vision globale. Nos dernières études ont porté sur la problématique de la fusion d'images pour permettre la détection des nuages de poussière dans la zone sahélienne d'une part, et pour la segmentation et la classification du couvert végétal dans les images de télédétection, d'autre part. Dans cette dernière étude, une fusion de sur-segmentations suivie d'une fusion de capteurs multispectral et panchromatique a permis d'obtenir des résultats pertinents et très intéressants pour des applications de classification du couvert végétal. Pour consolider notre expertise dans cette démarche, une thèse en cotutelle intitulée "Fusion d'images en télédétection satellitaire" débutera avant la fin de l'année 2009. Elle viendra ainsi enrichir les travaux déjà réalisés et permettre la mise en œuvre d'une chaîne de traitement et de fusion pour l'aide à l'interprétation.

Pour finir, on peut observer que d'un point de vue général, le développement sans cesse croissant de la technologie, nous oblige à adapter nos solutions pour traiter le flux d'information qui nous envahit chaque jour. Le besoin d'outils d'aide à l'interprétation se fait sentir et induit la nécessité de solutions facilitant à la fois l'exploration de grandes masses de données et permettant des prises de décisions rapides pour l'utilisateur. Nos études s'orienteront donc vers la convergence des deux aspects développés, à savoir la **description** et la **fusion**, pour aboutir à un schéma global les alliant pour une indexation d'images par le contenu. La fusion des descripteurs conduirait à une description sémantique. Ainsi, après une analyse fournissant des descripteurs symboliques de base (couleur, textures, formes...), une fusion de ces derniers avec des informations de contexte (context-aware) permettrait d'approcher une description sémantique.

X - Sigles utilisés

ADEME

Agence Départementale pour l'Economie et la Maîtrise d'Energie

AGER

Assistance à la Gestion de l'Environnement par Radar

ALER

ALlocataire d'Enseignement et de Recherche

AQL

Alliance Qualité Logiciel (actuel Silicomp-AQL)

ATER

Attaché Temporaire d'Enseignement et de Recherche

Bull's eye performance (BEP)

Dans un système de recherche d'information, c'est le rappel après récupération du double du nombre de pertinents (vrais positifs) correspondant à la requête.

CBI

Collaboration Bretagne Image : structure regroupant des industriels, des laboratoires universitaires et des centres de recherche.

CBIR

Content-Based Image Retrieval (CBIR) : c'est une application logicielle permettant de rechercher une image dans une base d'images en s'appuyant sur le contenu.

CBVIR

Content-Based Visual Information Retrieval (CBVIR) *voir* CBIR

CELA_r	Centre d'Electronique de L'Armement.
CPI	Carte de Points d'Intersection
CSS	Curvature Scale Space : c'est la carte des points d'inflexion du contour en fonction de l'échelle.
DEA	Diplôme d'Etudes Approfondies (actuel Master recherche)
DEM	Digital Elevation Model (<i>voir</i> MNT ₂)
DRI	Détection Reconnaissance Identification
DSP	Digital Signal Processor (Processeur de traitement numérique du signal)
d-TASS	differential-Turning Angle Scale Space : la carte d-TASS consiste en l'ensemble des points essentiels : passages par zéro (γ -points), minima (α -points) et maxima (β -points) de l'angle tangentiel ou angle de braquage (ou turning angle) d'une courbe.
EDP	Equations aux Dérivées Partielles
EII	département (de l'INSA de Rennes) Electronique et Informatique Industrielle
ERS	European remote sensing satellite
fBm	fractional Brownian motion (mouvement brownien fractionnaire)
FFT	Fast Fourier Transform ; transformée de Fourier rapide
GE	département (de l'INSA de Rennes) Génie Electrique (<i>voir</i> EII)
GMA	département (de l'INSA de Rennes) Génie Mécanique et Automatique

GP	département (de l'INSA de Rennes) Génie Physique (<i>voir</i> MNT ₁)
GUI	Graphical User's Interface
IPM	Intersection Points Map (<i>voir</i> CPI)
IPSIS	Ingénierie Pour Signaux et Systèmes
LPE	Ligne de Partage des Eaux
MNT₁	département (de l'INSA de Rennes) Matériaux et NanoTechnologie
MNT₂	Modèle Numérique de Terrain (ou DEM) : c'est une représentation de la topographie d'une zone terrestre.
MPEG	Moving Picture Experts Group : groupe de travail chargé du développement de normes internationales pour la compression, la décompression, le traitement et le codage de la vidéo, de l'audio et de leur combinaison.
MSGPR	Multi-Scale curve smoothing for Generalised Pattern Recognition : une approche d'analyse multiéchelle basée sur la carte des points d'intersection (CPI)
NTE	Nouvelles Technologies Educatives
PFE	Projet de Fin d'Etudes
Précision	Dans un système de recherche d'information et pour une requête donnée, c'est le rapport entre le nombre de pertinents récupérés sur le nombre d'éléments récupérés de la base.
QBIC	Query By Image Content : une solution CBIR proposée par IBM
QBE	Query By Example : langage de requête par l'exemple proposé par IBM

RdF	Reconnaissance des formes
Rappel	Dans un système de recherche d'information et pour une requête donnée, c'est le rapport entre le nombre de pertinents récupérés sur le nombre de pertinents dans la base.
ROS	Radar à Ouverture Synthétique ou RSO pour Radar à Synthèse d'Ouverture (voir SAR)
RRM	Rapport des Radiométries Moyennes, c'est le rapport R de la radiométrie moyenne des sous-fenêtres d'analyse ; il est normalisé en r compris entre 0 et 1 en considérant $r = \min(R, 1/R)$
SAR	Synthetic Aperture Radar (<i>voir</i> ROS)
SIROS	Simulateur d'Images Radar à Ouverture Synthétique
SLC	Single Look Complex
SPOT	satellite pour l'observation de la terre
TELEDITS	TELEopération de DSP par Internet pour Traitement du Signal : outil logiciel pour l'apprentissage et le partage des DSP sur internet.
VOAD	Vertical Aerosol Optical Depth : images produites à partir du canal visible de METEOSAT pour aider à la detection de la poussière au-dessus du Sahara

XI - Bibliographie

XI.1 - Références bibliographiques

- [Ahmad98] Dr. Ejaz Ahmad (Deputy Director General of WWF-Pakistan), *The root causes of biodiversity loss in the Indus delta*, Analysis of a study conducted by WWF-Pakistan in 1998 with financial support of DANIDA, DGIS, BMZ, GEF and WWF-Sweden.
- [Baraldi95] Baraldi A., Parmiggiani F., *An alternative form of the Lee filter for speckle suppression in SAR images*, Graphical Models and Image Processing, vol. 57, n° 1, pp. 75-78, 1995
- [Barenholtz03] Barenholtz E., E. H. Cohen, J. Feldman and M. Singh, *Detection of change in shape: an advantage for concavities*, Elsevier, Cognition 89, pp. 1–9, 2003
- [Barnsley88] Barnsley M. F., *Fractals Everywhere*, Academic Press, Inc., 1250 Sixth Avenue, San Diego, 1988
- [Bentoutou06] Bentoutou Y., Taleb N., Kpalma K. and Ronsin J., *A Feature-Based Approach to Automated Registration of Remotely Sensed Images*, 2nd IEEE International Conference on information & Communication Technologies: from theory to applications (ICTTA'06), Damascus, Syria, April 24-28 2006
- [Bentoutou05] Bentoutou Y., Taleb N., Kpalma K. and Ronsin J., *An Automatic Image Registration for Applications in Remote Sensing*, IEEE Transactions on Geoscience and Remote Sensing, Vol. 43, No. 9, pp 2127-2137, September 2005
- [Bentoutou03] Bentoutou Y., Taleb N., Kpalma K. and Ronsin J., *An invariant approach for automated image registration of multi-temporal remotely sensed imagery*, the conference on Remote Sensing for Environmental Monitoring, GIS Applications, and Geology III, part of

- the International Symposium on Remote Sensing, Barcelona, Spain, 8-12 September 2003
- [Berroir98] Berroir, J.-P. and Béréziat, D., *Cloud Cover Estimation for Air Quality Models using Meteorological Sensors*, Systems Analysis Modeling Simulation, 1998.
- [Bertrand77] Bertrand J.J., *Action des poussières sub-sahariennes sur le pouvoir glaçogène de l'air en Afrique de l'Ouest*, Thèse de Doctorat d'Etat, Université de Clermont-Ferrand, 230 pp, 1977.
- [Bloch96] Bloch, I. *Information combination operators for data fusion: a comparative review with classification*, IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans, vol. 26, pp. 52-67, January 1996
- [Bloch94] Bloch I., Maître H., *Fusion de données en traitement d'images : modèles d'information et décision*, Traitement du Signal, vol. 11, n° 6, pp. 435-446, décembre 1994
- [Brooks00] Brooks, N., *Dust-climate interactions in the Sahel-Sahara zone of northern Africa, with particular reference to late twentieth century Sahelian drought*. PhD Thesis, Climatic Research Unit, University of East Anglia, Norwich, 2000
- [Brown92] Brown L.G., *A survey of image registration techniques*, ACM Computing Surveys, vol. 14, n° 4, pp. 325-376, December 1992
- [Coster85] Coster M., Chermant J.-L., *Précis d'analyse d'images*, Editions du C.N.R.S., 15, quai A. France, Paris, 1985
- [Couturier02] Couturier C., Kpalma K., Ronsin J., *DSP Teleoperation for Digital Signal Processing Teaching and Learning via Internet*, Fourteenth International conference on DSP (DSP2002), Santorini, Greece, 1-3 July 2002
- [Denos93] Denos Y.L., Matteini V., *Review on structure detection and speckle filtering on ERS-1 images*, EARSel Advances in Remote Sensing, Vol. 2, No. 2 - VI, pp 52-65, 1993
- [Descombes95] Descombes X., *A fission and fusion Markovian approach for multi-channel segmentation*, IEEE International Geoscience and Remote Sensing Symposium Proceedings, pp. 124-127, July 1995
- [Diop06] Diop O., Kpalma K. and Ronsin J., *Textural approach for the detection of the dust clouds from infrared satellite*, Part of the SPIE Symposium on Asia-Pacific Remote Sensing, 13-17 November 2006, Goa, India. (à paraître)
- [Dhonneur74] Dhonneur G., *Nouvelle approche des réalités météorologiques de l'Afrique Occidentale et Centrale*, Thèse de Doctorat Ingénieur, UCAD de Dakar, Tome1 358 pp, Tome2 472 pp, 1974.
- [Feldman05] Feldman J. and M. Singh, *THEORETICAL NOTE: Information Along Contours and Object Boundaries*, Psychological Review, The

- American Psychological Association, Vol. 112, No. 1, pp. 243–252, 2005
- [Garba92] Garba A., *Approche des modifications des lignes de grain en zone côtière Ouest-Africaine*, Thèse de Doctorat Ingénieur, UCAD de Dakar, 163 pp, 1992.
- [Goshtasby88b] Goshtasby A., *Image registration by local approximation methods*, Image and Computing Vision, vol. 6, n° 4, pp. 255-261, November 1988
- [Goshtasby88a] Goshtasby A., *Registration of image with geometric distortions*, IEEE, Transactions on Geoscience and Remote Sensing, vol. 26, n° 1, pp. 60-64, January 1988
- [Haese-Coat94] Haese-Coat V., Kpalma K., Ronsin J., *Apport comparatif de la texture en segmentation d'images SPOT*, Revue Photo-Interprétation, nov. 1994, Editions ESKA
- [Hochchild00] Hochchild V., and Herlin I., and Moretti S., and Ranchin T., and Staudenraush H., *Application of Remote Sensing to the Development of an Integrated Water Resources Management System (IWRMS)*, 28th International Symposium on remote sensing of environment, Cape Town, South Africa, March 2000.
- [Jankoviak92] Jankoviak I., Touré D., *Les climats subtropicaux et leur évolution: de l'observation spatiale à la modélisation*, Aérosols désertiques : Sources, transports, impacts par télédétection La-Londe-Les-Maures, Septembre 1992.
- [Kalluri04] Kalluri S., Gilruth P., *Moving from Research to Applications: NASA's EOS Data for Decision Makers*, The 84th American Meteorological Society (AMS) Annual Meeting, Second Annual User's Conference , 14 January 2004
- [Kaneko89] Kaneko H., *A Generalized Fractal Dimension and Its Application to Texture Analysis*, IEEE, 1989, pp 1711-1714
- [Kpalma08] Kpalma K., Yang M., Ronsin J., Planar shapes descriptors based on the turning angle scalogram, 5th International Conference on Image Analysis and Recognition (ICIAR 2008), Povia de Varzim, Portugal, June 25-27, 2008. Proceedings, Lecture Notes in Computer Science, Vol. 5112, Springer 2008, ISBN 978-3-540-69811-1
- [Kpalma07c] Kpalma K., Ronsin J., *An overview of advances of pattern recognition systems in computer vision*; in: "Vision systems : segmentation and pattern recognition", Chapter 10; Obinata and Dutta (Eds.), I-Tech Education and Publishing, Vienna, 2007
- [Kpalma07b] Kpalma K., Ronsin J., *A Novel Multi-Scale Representation for 2-D Shapes*, 4th International Conference on Image Analysis and Recognition (ICIAR 2007), Montreal, Canada, August 22-24, 2007, Proceedings. Lecture Notes in Computer Science 4633 Springer 2007, ISBN 978-3-540-74258-6

- [Kpalma07a] Kpalma K., Ronsin J., *Turning Angle Based Representation for Planar Objects*, Electronics Letters, 10 May 2007, Volume 43, Issue 10, p. 561-563
- [Kpalma06b] Kpalma K., Ronsin J., Description des Formes Planaires par Lissage Progressif de Contours Fermés, 8e Colloque Africain sur la Recherche en Informatique (CARI), 6-8 novembre 2006 à Cotonou, Bénin
- [Kpalma06a] K. Kpalma, J. Ronsin, *Multiscale contour description for pattern recognition*, Pattern Recognition Letters, Vol.27, No.13, pp 1545-1559, 1 October 2006
- [Kpalma05] Kidiyo Kpalma and Joseph Ronsin, Contours Smoothing for Non-Occluded Planar Shapes Description, International Conference on Image Processing (ICIP 2005), Genova, Italy, September 11-14, 2005
- [Kpalma03d] Kpalma K., Ronsin J., *Planar Curve Smoothing for Pattern Recognition*, Third International Symposium on Multispectral Image Processing and Pattern Recognition (MIPPR'03), 20-22 October 2003, Beijing, China, in Proc. SPIE the International Society for Optical Engineering; 2003, Issue 5286, Part 1, pp. 367-370
- [Kpalma03c] Kpalma K., *TELEDITS : enseignement de DSP à distance*, Séminaire Quels DSP pour les applications de télécommunications et images ?, organisé par INSA de Rennes, Texas Instruments et ATEME, 07 octobre 2003
- [Kpalma03b] Kpalma K., Haese-Coat V., *Traitement numérique du signal : Théorie et applications*, Collection Technosup, Editions Ellipses, Septembre 2003, ISBN 272981728X
- [Kpalma03a] Kpalma K., Ronsin J., *A Multi-Scale curve smoothing for Generalised Pattern Recognition (MSGPR)*, Seventh International Symposium on Signal Processing and its Applications (ISSPA), 1-4 July 2003, Paris, FRANCE
- [Kpalma01] Kpalma K., Ronsin J., Adam O., *Rapport Final (RP2)*, Rapport de convention "Algorithmies pour l'aide à la détection, la reconnaissance et l'identification (DRI) des cibles militaires dans les données de l'observation. Poste 2 - Explosion / Fusion algorithmique - Marché CELAR n° 99 42 165 00 470 35 15", février 2001
- [Kpalma96] Kpalma K., Ronsin J., *Génération de la trajectoire de meule pour le contrôle d'une machine à commandes numériques dans un système de d'affûtage de scies à bois*, Rapport de convention INSA-ALLIGATOR (étude et développement), mars 1996
- [Kpalma94] Kpalma K., *Caractérisation de textures par l'anisotropie de la dimension fractale*, 2e Colloque Africain sur la Recherche en Informatique (CARI), octobre 1994 à Ouagadougou, Burkina Faso.
- [Kpalma93b] Kpalma K., Haese-Coat V., Ronsin J., *A Hierarchical Multiresolution Texture Image Segmentation*, SPIE-Visual Communication and Image Processing, November 1993, Boston, USA

- [Kpalma93a] Kpalma K., Haese-Coat V., *Une Nouvelle Méthode de Segmentation d'Images de Textures Multirésolution*, Colloque de GRETSI, septembre 1993 à Juan-les-Pins, France
- [Kpalma92c] Kpalma K., Ronsin J., *Modélisation et numérisation de profils de denture de scie*, rapport final de la convention INSA-ALLIGATOR, Réf : CR368, INSA-Rennes, juin 92.
- [Kpalma92b] Kpalma K., Ronsin J., *Modélisation et numérisation de profils de denture de scie*, rapport intermédiaire de la convention INSA-ALLIGATOR, Réf : CR368, INSA-Rennes, mars 92.
- [Kpalma92a] KPALMA K., *Analyse Fractale de Textures Naturelles dans un Contexte Multirésolution : Application à la Segmentation d'Images Multirésolution*, Thèse de Doctorat de l'INSA de Rennes, 12 février 1992.
- [Kpalma91b] Kpalma K., Bruno A., Haese-Coat V., *Natural Texture Analysis in Multiscale Context Using Fractal Dimension*, Proc. SPIE, Visual Commun. Image Processing, November 1991, Boston, USA
- [Kpalma91a] Kpalma K., Haese-Coat V., Bruno A., *Analyse Pyramidale de textures Naturelles à Travers la Dimension Fractale*, 13^e colloque GRETSI, 16-20 septembre 1991, Juan-les-Pins, France, pp 1053-1056
- [Lee95] Lee J.S., *A simple speckle smoothing algorithm for synthetic aperture radar images*, IEEE Transactions on Systems, Man and Cybernetics, vol. 13, No. 1, pp 85-89, January/February 1983
- [Legrand94] Legrand M., N'Doume C. and Jankowiak I., *Satellite-derived climatology of the Saharan aerosol, Passive Infrared remote sensing of clouds and the atmosphere II*, SPIE 2309, pp127-135, (1994)
- [Lindeberg94] Lindeberg T., *Scale-Space Theory in Computer Vision*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1994.
- [Lopès93] Lopès A., Nezri E., Touzi R., Laur H., *Structure detection and statistical adaptive speckle filtering in SAR images*, International Journal of Remote Sensing, vol. 14, n° 9, pp. 1735-1758, 1993
- [Le Hégarat97] Le Hégarat-Masclé S., Bloch I., Vidal-Madjar D., *Application of Dempster-Shafer evidence theory to unsupervised classification in multi-sources remote sensing*, IEEE Transactions on Geoscience and Remote Sensing, vol. 35, n° 4, pp. 1018-1031, July 1997
- [Matusiak99] Matusiak S., *Description Invariante et locale des formes planes Application à l'indexation d'une base d'images*, Thèse de Doctorat, soutenue à l'Université de Valenciennes et du Hainaut Cambresis, 17 septembre 1999 à l'ENIC/INT (Lille)
- [Matusiak98] Matusiak S., Daoudi M., *Planar Closed Contour Representation by Invariant Under a General Affine Transformation*, IEEE International Conference on System, Man and Cybernetics (IEEE-SMC'98), pp. 3251-3256, October 11-14, 1998, Hyatt Regency La Jolla, San Diego, California, USA

- [Memon05] Altaf A. Memon, Ph.D., M.ASCE, *Devastation of the Indus River Delta*, Proceedings, World Water & Environmental Resources Congress 2005, American Society of Civil Engineers, Environmental and Water Resources Institute, Anchorage, Alaska, May 14-19, 2005. (Note – Expanded version)
- [Mokhtarian96] Mokhtarian F., Abasi S., Kittler J., *Efficient and Robust Retrieval by Shape Content through Curvature Scale Space*, in Proc. International Workshop on Image Databases and MultiMedia Search, 1996, 35-42, Amsterdam, The Netherlands
- [Mokhtarian92] Mokhtarian F., Mackworth A. K., *A Theory of Multiscale, Curvature-Based Shape Representation for Planar Curves*, in IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-14, N° 8, August 1992
- [MPEG-7,1] MPEG Requirements Group, *MPEG-7: Overview*, ISO/IEC JTC1/SC29/WG11N5525, Pattaya, March 2003
- [MPEG-7,2] MPEG Requirements Group, *MPEG-7 Requirements Document*, Doc. ISO/MPEG N3548, MPEG Beijing Meeting, July 2000.
- [Nezri92] Nezri E., *Restauration de réflectivité radar pour l'utilisation conjointe des images radar et optiques en télédétection*, Thèse de Doctorat, Université Paul Sabatier, Toulouse, n° 1236, juillet 1992
- [N'Tchayi 94] N'Tchayi, G. M., Bertrand, J., Legrand M. and Baudet, J. Temporal and spatial variations of the atmospheric dust loading throughout West Africa over the last thirty years, *Annales Geophysicae* 12, 265-273, 1994.
- [Ogor97] Ogor B., *Etude comparative de méthodes de filtrage de speckle et de segmentation en imagerie radar à synthèse d'ouverture : application à la région des polders du Mont-Saint-Michel*, Thèse de Doctorat, INSA de Rennes, décembre 1997
- [Ogor96] Ogor B., Haese-Coat, V., Ronsin J., *SAR image segmentation by mathematical morphology and texture analysis*, IEEE International Geoscience and Remote Sensing Symposium Proceedings, vol. 1, pp. 717-719, May 1996
- [Pentland84] Pentland A. P., *Fractal-Based Description of Natural Scenes*, IEEE PAMI, vol. 6, No. 6, November 1984, pp 661-674
- [Pereira99] Pereira F., *MPEG-4 : Why, What, How and When?*, Image Communication Journal, vol. 15, n° 4-5, December 1999.
- [Safa89] Safa F., Flouzat G., *Speckle removal on radar imagery based on mathematical morphology*, Signal Processing, Vol. 16, pp 319-333, 1989
- [Saggi95] Saggi M., Boucher J.M., Béné G., *Hierarchical filtering of SAR images*, Proceedings of IGARSS'95, Florence (Italy), Vol. 2, pp 898-900, July 1995

- [Sebastian04] Sebastian T., Klein P., and Kimia B., *Recognition of shapes by editing their shock graphs*, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26(5), pp. 550_571, 2004.
- [Sossa00] Sossa H., *Object Recognition, Summer School on Image and Robotics*, INRIA Rhône-Alpes, June 26 - July 7, 2000
- [Thépaut00c] Thépaut O., Kpalma K., Ronsin J., *Automatic registration of ERS and SPOT multisensor images in a data fusion context*, Elsevier's Forest Ecology and Management, Vol. 128, pp 93-100, 2000
- [Thépaut00b] Thépaut O., Kpalma K., Ronsin J., *Spécification de la maquette de démonstration*, Rapport de convention "Algorithmies pour l'aide à la détection, la reconnaissance et l'identification (DRI) des cibles militaires dans les données de l'observation - Marché CELAR n° 99 42 165 00 470 35 15", novembre 2000
- [Thépaut00a] Thépaut O., Kpalma K., Ronsin J., *Rapport Technique du Poste 2 (RTP2)*, Rapport de convention "Algorithmies pour l'aide à la détection, la reconnaissance et l'identification (DRI) des cibles militaires dans les données de l'observation - Marché CELAR n° 99 42 165 00 470 35 15", septembre 2000
- [Thépaut99] Thépaut O., Kpalma K., Ronsin J., *Recalage d'images de télédétection multi-source appliqué au filtrage multirate et à la fusion de segmentation*, Colloque International La Télédétection Optique et Radar et la Géomatique pour la Gestion des Problèmes Environnementaux, 67e Congrès de l'ACFAS, 18 pages, Ottawa, Ontario, Canada, mai 1999
- [Thépaut98d] Thépaut O., Kpalma K., Ronsin J., *Géocorrection et Recalage d'images ERS et SPOT dans un contexte de fusion de données*, 4e Colloque Africain sur la Recherche en Informatique (CARI), p. 891, octobre 1998 à Dakar, Sénégal.
- [Thépaut98c] Thépaut O., Kpalma K., Ronsin J., *Multichannel remote sensing images registration in a consensual segmentation scheme*, The European Symposium on Remote Sensing, Conference on Image and Signal Processing for Remote Sensing IV, Vol. 3500, 21-27 Sept. 1998, Barcelona, Spain
- [Thépaut98b] Thépaut O., Kpalma K., Ronsin J., *ERS SAR and SPOT Images Geocorrection and Registration in a context of Data Fusion*, Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, pp. 1040-1042, July 1998, Seattle, Washington, USA.
- [Thépaut98a] Thépaut O., Kpalma K., Ronsin J., *Full Automatic Method of ERS and SPOT Images Registration in a Context of Data Fusion*, Proceedings of the First International Conference on Geospatial Information in Agriculture and Forestry, vol. 1, pp. 654-661, June 1998, Lake Buena Vista, Florida, USA
- [Thépaut96] Thépaut, O., *Etude comparative d'algorithmes de segmentation en imagerie SAR : application à l'extraction de parcelles agricoles*,

- Mémoire de stage de DEA en Electronique de l'INSA de Rennes, 58 pages, juillet 1996
- [Touzi88b] Touzi R., Lopès A. and Bousquet, P., *A statistical and geometrical edge detector for SAR images*, IEEE Transactions on Geoscience and Remote Sensing, vol. 26, n° 6, pp. 764-773, November 1988
- [Touzi88a] Touzi R., *Analyse d'images radar en télédétection : amélioration radiométrique, filtrage du speckle et détection de contours*, Thèse de Doctorat, Université Paul Sabatier, Toulouse, n° 258, mars 1988
- [Tupin96] Tupin, F., Gouinaud, C., Maître, H., Crettez, J.P., Nicolas, J.M., *Détection de structures linéaires sur des images ROS - Linear feature detection on SAR images*, Traitement du Signal, vol. 13, n° 6, pp. 635-650, spécial 1996
- [Vapillon98] A. Vapillon, B. Collin et A. Montanvert, *Analyzing and Filtering Contour Deformation*, International Conference on Image Processing (ICIP), Chicago, Illinois, USA, October 4-7, 1998
- [Vapillon97] A. Vapillon, B. Collin and A. Montanvert, *Appariement de contours 2D par analyse multirésolution hiérarchique de la déformation*, proceedings of GRETSI'97, Grenoble, France, Sept. 1997.
- [Yang07b] Yang M., Kpalma K., Ronsin J., *Affine invariance contour descriptor based on iso-area normalization*, IEE Electronics Letters, Vol.43, No7, pp.379-380, (2007)
- [Yang07a] Yang M., Kpalma K. and Ronsin J., *A Shape Descriptor based on CSS*, Document Recognition and Retrieval XIV Part of the IS&T/SPIE International Symposium on Electronic Imaging , 28 January - 1 February 2007, San Jose Marriott and San Jose Convention Center San Jose, California, USA
- [Zamperoni93] Zamperoni, P., *An adaptive rank-order filter for image enhancement depending on a measure of the local spatial order*, 14^{ème} Colloque GRETSI, Juan-Les-Pins, vol. 1, pp 543-547, 1993
- [Zibreira01] Zibreira C., Pereira F., *A Study of Similarity Measures for a Turning Angles-based Shape Descriptor*, 3rd Conference on Telecommunications, conftele 2001, Figueira da Foz, 23-24 April 2001

XI.2 - Références webographiques

- [1] SQUID : <http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html> : ce site propose une solution de recherche d'images basée sur la carte CSS. Il offre, également, une base d'images à télécharger et une démonstration de recherche de formes SQUID (Shape Queries Using Image Databases) (vérifier le lien)
- [2] Google : <http://www.google.fr/imghp> : moteur de recherche d'images par mots-clé
- [3] IMEDIA : <http://www.inria.fr/recherche/equipes/imedia.fr.html> : IMEDIA est un projet de recherche sur l'indexation, la navigation et la recherche d'images dans une base d'images
- [4] Photobook : <http://vismod.media.mit.edu/vismod/demos/photobook/> : moteur de recherche d'images basé contenu.
- [5] <http://amazon.ece.utexas.edu/~qasim/research.htm> : ce site propose l'outil CIRES (Content Based Image REtrieval System) utilisant une approche combinant les principes de la perception de bas niveau et de haut niveau.
- [6] VIPER : <http://viper.unige.ch/research/> : groupe de travail sur le thème du traitement de l'information visuelle pour la recherche améliorée (Visual Information Processing for Enhanced Retrieval)
- [7] PIRIA : http://www-list.cea.fr/fr/programmes/systemes_interactifs/labo_lic2m/piria/w3/pirianet.php
- [8] SIMPLIcity : http://wang14.ist.psu.edu/cgi-bin/zwang/regionsearch_show.cgi
- [9] Corbis : <http://corbis.demo.ltutech.com/en/demos/corbis/>
- [10] FlickrR : <http://www.flickr.com/>
- [11] Retrievr : <http://labs.systemone.at/retrievr/>
- [12] Xcavator : <http://www.xcavator.net>

XII - Index

A

ADEME 29, 43, 52, 111
Afrique 44, 116
AGER 20, 53, 111
ALER 15, 41, 111
analyse
 de textures 16, 19, 129
 des détails 74
 multiéchelle 11, 28, 37, 59, 61, 94, 96, 99, 109
 multi-échelle 101, 189, 205
 multirésolution 122
angle
 tangential 97, 98, 100
AQL 43, 111
ATER 15, 41, 111
attributs 19, 20, 28, 60
 caractéristiques 60, 68, 71, 76
AUC 103

B

BEP 103
Bull's eye 103, 111

C

carte
 CPI 68
 des points d'intersection 68
 d-TASS 100, 102
CBI 53, 111
CBIR 59, 107, 111, 113
CBVIR 59, 111
CELA_r 112
chat dormant 98
classification 20, 21, 24, 25, 49
concave 97
contour 28, 97
 fermé 28, 61
 filtré 61, 70, 94, 99

 lissé 62, 66, 67, 68, 100
 original 62, 67, 68, 100
contours
 filtrés 28
convexe 97
courbure 61, 97, 98, 99
CPI 28, 62, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 86, 87,
 90, 91, 94, 96, 97, 107, 109, 112, 113
CSS 28, 49, 60, 61, 65, 85, 86, 87, 88, 89, 90, 94, 99,
 106, 112, 122, 125
curvature 61

D

DEA 15, 39, 40, 112, 122
DEM 112, 113
DRI 22, 23, 43, 52, 112, 118, 121
DSP 16, 29, 33, 40, 41, 50, 52, 53, 112, 116, 118
d-TASS 99, 100, 101, 102, 104, 106, 107, 109, 112

E

EDP 112
EII v, vi, 15, 34, 41, 42, 44, 53, 54, 112
ERS 21, 22, 38, 48, 50, 51, 112, 116, 121, 122, 151

F

fBm 20, 112
FFT 112
filtrage
 progressif 60, 109
filtrage
 gaussien 61
 passe-bas 28, 62, 65, 99
filtre
 gaussien 65, 68, 69, 70, 100
FOOS 25
formes planaires 28, 61, 94, 118
fractale 20, 46, 47, 51
fréquences
 basses 70

hautes 70, 74, 75
fusion
 de données 22, 51, 121
 d'images 16, 21, 22, 37, 43, 96, 107, 110
 pour la visualisation 23, 108

G

GE 112
GMA 113
GP 113
GUI 40, 113

I

IETR 16, 43
INSA vi, 15, 16, 41, 44
IPM 68, 87, 113
IPSIS 22, 43, 113

K

k-means 25
kNN 26

L

lissage
 progressif 99
lissage 28, 60, 62, 65, 69, 94, 99, 100
 progressif 28
LPE 113

M

Météosat 24, 25
MNT 21, 41, 54
MPEG 88, 89, 113, 120, 121
MPEG7 60, 88, 89, 120
MSAVI2 26
MSGPR 28, 50, 60, 61, 62, 85, 89, 96, 107, 113, 118
multicanal 22
multicapteur 129, 151
multicapteurs 16, 21
multidate 51, 121
multidates 21
multirésolution 20, 46, 47, 131
multi-scale 28
multisource 51, 119, 121
multispectrales 21

N

NDVI 26
NGRDI 26
Normalisation 30, 62, 71, 97
NTE 33, 44, 113
nuages de poussière 24, 39, 49, 110

O

objets planaires 11, 59

P

panchromatique 26
PFE v, 37, 39, 40, 113
points
 d'intersection 28, 61, 62, 66, 68, 69, 70, 94
 essentiels 98, 100, 102, 104, 106
points essentiels 98, 101
Précision 103, 113

Q

QBIC 59, 113
Quickbird 25

R

radar 22, 38, 39, 107, 110, 119, 120, 121, 122
Rappel 103, 114
RdF 11, 96, 114
recalage d'images 22, 44, 107, 129
ROS 21, 114, 122
RRM 22, 114

S

SAR 21, 22, 39, 43, 50, 110, 114, 115, 119, 120, 121, 122
SAVI 26
scale 61
scalogramme 99, 101
segmentation 16, 20, 21, 22, 38, 39, 46, 47, 48, 50, 51, 107, 110, 116, 117, 120, 121, 122, 129, 141
signature 20
similarité 60, 78, 79, 80, 83, 84, 85, 86, 87, 89, 90, 91, 93
SIROS 22, 43, 114
SLC 22, 23, 114
space 61
SPOT 20, 21, 22, 38, 48, 50, 51, 114, 117, 121, 122, 141, 151
sur-segmentation 110, 173

T

télédétection 16, 17, 43, 44, 51, 60, 121
TELEDITS 33, 34, 35, 40, 53, 114
texture
 attributs 25
 images de 131
transformation
 affine 181
transvection 102, 103
turning angle 112

V

VOAD 25, 114

Troisième partie

Recueil des travaux

Cette partie présente huit articles issus de mes travaux de recherche. Ces articles retracent l'évolution et marquent la continuité dans mes travaux de recherche depuis l'analyse de textures à la segmentation d'images de télédétection en passant par le recalage d'images multicapteur. Parallèlement à cet aspect d'analyse d'images orientés vers la segmentation et la classification, ces travaux illustrent nos activités de recherche en analyse pour la description des formes et pour l'indexation d'images.

Pour chaque article, on trouvera, en en-tête, la liste des auteurs, le titre, la revue ou la conférence suivi d'un petit résumé en français.

XIII - Analyse multirésolution d'images de textures

Kidiyo Kpalma, Véronique Haese-Coat et Joseph Ronsin

A Hierarchical Multiresolution Texture Image Segmentation

Proc. SPIE Vol. 2094, p. 870-877, Visual Communications and Image Processing '93

SPIE-The International Society for Optical Engineering

Novembre 1993, pp 870-877 (8 pages)

Résumé

La segmentation d'images de textures est un problème qui a fait l'objet de nombreuses études. Diverses méthodes, comme celles utilisant plusieurs images de résolutions différentes (ou niveaux de pyramide) sont ainsi proposées. Dans des travaux antérieurs, nous avons présenté une méthode d'analyse multirésolution aboutissant à un algorithme de segmentation (ASM) qui s'est révélé performant mais complexe dans sa mise en oeuvre. Pour simplifier le processus, nous avons proposé un algorithme hiérarchique (ASH) qui segmente les différentes résolutions, l'une après l'autre. Par cette simplification, le système s'est alors trouvé pénalisé par une perte de performances (5%) mais rapporte un gain en termes de temps de calculs soit une réduction moyenne 50%.

A HIERARCHICAL MULTIREOLUTION TEXTURE IMAGE SEGMENTATION

K. KPALMA, V. HAESE-COAT and J. RONSIN

Laboratoire d'Automatique, Equipe Image
Institut National des Sciences Appliquées
35043 RENNES Cedex (FRANCE)

1. INTRODUCTION

In this work we present a method for multiresolution texture image segmentation via supervised Bayesian classification. This method takes into account a function called *classification index CI* (or Indice de Classification IC³). This function measures the accuracy rate of a pixel classification after it has been classified. The CI measure is based on the location of the pixel attribute vector in the observation space relatively to the classes to which must be classified.

In some former papers^{1,2,3}, we proposed the ASM method (ASM : Algorithme de Segmentation Multiresolution³) that combines information extracted from different pyramid levels (multiresolution images) to build the global attribute vector for Bayesian classification purpose. The method has then been found to be efficient but large time consuming in the segmentation process. It is known that the number of classification operations (matrix operations) increases like n^2 ; n being the dimension of the covariance matrix. Since we use many resolutions of the image under study, the dimension of the global attribute vector is then multiplied by the number of the multiresolution images that are taken into account. The time needed in the classification process then becomes very large. In this method, if r is the number of different resolution images and if one extracts p features at each resolution image, then the attribute vector is $n=rp$ large. Obviously, the number of matrix operations is then r^2 times greater than it would be for only one resolution image. To reduce time used in the classification process, we propose a new method (ASH : Algorithme de Segmentation Hiérarchique³) that segments hierarchically the multiresolution images.

2. ASM METHOD

To contribute to texture segmentation that is an important problem in image processing, we proposed a method^{1,2,3} that takes into account different resolutions of the same image. This allows segmentation of textured images in multiresolution context. We showed in these works that segmentation methods using multiresolution images were more efficient than those using a unique resolution image. We could obtain about 99% rate of good segmentation result on some textures by taking into account two or three pyramid levels. This method called ASM (for Multiresolution Segmentation Algorithm) is then found to be efficient. The principle of the ASM is to extract features at different resolution images and concatenate them to build the global attribute vector for Bayesian classification purpose. By doing so, the attribute vector is $n=rp$ long and as it has been said earlier, the computation time increases as r becomes greater.

Since the method is found to be efficient, it is a need to find another multiresolution method with the same performances but much faster. We then propose the ASH method for hierarchical segmentation of the multiresolution images.

3. ASH METHOD

To reduce the number of the ASM's classification operations, we propose to deal with resolution images one after another and to combine the segmentation results. The idea is to proceed by different steps corresponding to different resolution image segmentation. At each step, we classify pixels and then detect and reject the misclassified ones. At the next step, we classify only the pixels that were rejected at the previous step. This allows computation time reduction since the attribute vector is less dimensional ($\dim=p$) and we have just a fraction of the image to deal with from a step to another. To do so, we have to define misclassified pixels detector. The classification index function is then proposed to measure the accuracy of a pixel classification that will help misclassification detection.

3.1. CLASSIFICATION INDEX FUNCTION DEFINITION.

This function measures the confidence rate of a pixel classification after it has been classified. In a Bayesian classification process, it appears necessary, sometimes, to know how accurate the classification is. The CI function that takes into account the distances between the pixel under study and the centre of the classes provides us of this measure. To explain the principle of this function, consider a Bayesian classification problem with two classes ω_1 and ω_2 characterised with two features a_1 and a_2 . The attribute vector is then of length 2. Considering a system with two features simplifies the problem and allows easily to represent the classes in a two-dimensional observation space. By representing the individuals of both classes in the observation space, we can distinguish two clusters corresponding to each of the classes. The clusters are centred on their representatives \bar{x}_1 and \bar{x}_2 which are the mean vectors of the corresponding cluster. Figure 1 represents a 2D system of two classes.

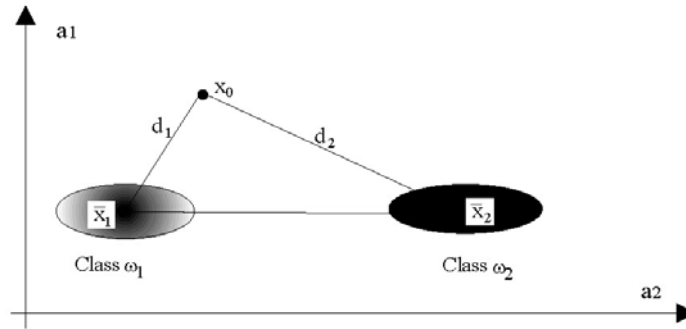


Fig. 1 : Representation of two classes in a 2D space

In a classical Bayesian classification, we use discriminant functions such as the Mahalanobis distances related to each class. For a class ω_s ($s=1,2$), the Mahalanobis distance is defined by

$$d_s(x) = (x - \bar{x}_s)^T \Sigma^{-1} (x - \bar{x}_s) \quad (1)$$

where \bar{x}_s is the mean vector of class ω_s and Σ_s the corresponding covariance matrix. The decision rule is then expressed by the expression below. An observed x is classified to class ω_s if it is nearer to \bar{x}_s than it is to \bar{x}_t ($t \neq s$).

$$\text{if } d_s(x) < d_t(x) \text{ (} s \neq t \text{) then decide } x \in \omega_s.$$

Commonly, we assume that the classes are different enough so that their clusters don't overlap. The decision border is then defined by a line lying between them. This border is a set of points for which different Mahalanobis distances are equal ($d_s = d_t$). It divides the observation space into two parts corresponding to each class. Each point then lies on the side of the border that correspond to the class at which it belongs.

In practice, the clusters overlap so that one can observe points belonging to both classes. On figure 2, E_s ($s=1,2$) is a set of points of class ω_s that lies outside the border line on the side of ω_t ($t \neq s$). This may occur when the features are not enough discriminant. The classical Bayesian classification process will classify E_2 into class ω_1 and E_1 into class ω_2 as it can be observed on figure 2. This produces misclassification errors corresponding to the overlapping areas E_1 and E_2 . It is then useful to define a third class ω_R that will contain misclassified points. To define this class we have to detect misclassification. To do so we need to have a measure to estimate the accuracy of a point classification. The classification index function is then used to get this measure. This function is defined by the distances between the observation and the centres of different classes (see figure 1). For an observation x , there are a minimum distance d_{\min} and a maximum distance d_{\max} between x and the mean vectors of both classes.

$$d_{\min} = \underset{s=1,2}{\text{MIN}}(d_s(x)) \text{ and } d_{\max} = \underset{s=1,2}{\text{MAX}}(d_s(x))$$

The classification index function is then defined by

$$CI(x) = 1 - \frac{d_{\min}}{d_{\max}} \quad (2)$$

This function varies from 0 to 1. The zero value corresponds to the case for which x is on the decision border ($d_1(x) = d_2(x)$). This is the most litigious case ; it is impossible to classify x without classification error. $CI(x) = 1$ ($d_{\min}=0$) corresponds to case for which x equals the mean vector of the nearest class. The classification of such a point is more accurate than the former one.

For class ω_R definition purpose, we need to fix a threshold CI_0 so that points with CI less than this value are to be rejected. Using a threshold allows us to reject points of E_1 and E_2 for further classification. In his study⁴, when Weber used the confidence rate to measure the classification accuracy, he chose the threshold arbitrarily. But as it can be seen on fig. 3.c, the mean value of the CI measure depends on the current class so it is a need to define threshold relatively to the class. It is a withdraw to choose a global arbitrary threshold without taking into account the mean values of different textures. If the threshold value is too high some whole classes may be rejected and if it is too low all points will be classified. In this case it appears useless to define a rejection class since many points are classified.

To overcome this drawback, we propose an automatic threshold computing by taking into account the configuration of different classes and relatively to the classes. On figure 2, \hat{x}_2 represents the point of class ω_2 which is closest to class ω_1 and in the same way, \hat{x}_1 is the point of class ω_1 which is closest to class ω_2 . These points have the greatest probability to cause classification errors so they are to be taken into account in the CI threshold definition.

Let δ_{12} be the distance between \hat{x}_2 and \bar{x}_1 observing distance of \hat{x}_2 to class ω_1 ($\delta_{12}=d_1(\hat{x}_2)$). If Δ_{22} is the distance between \hat{x}_2 and \bar{x}_2 relatively to class ω_2 ($\Delta_{22}=d_2(\hat{x}_2)$) then the classification index threshold for class ω_1 can be defined by

$$CI_0^1 = 1 - \frac{\delta_{12}}{\Delta_{22}} \quad (3)$$

Similarly the CI threshold related to class ω_2 is defined by

$$CI_0^2 = 1 - \frac{\delta_{21}}{\Delta_{11}} \quad (4)$$

Equations (3) and (4) allows to define the rejection points relatively to their class. Since we have to deal with more than two classes, we must define a similar method for multiclass case.

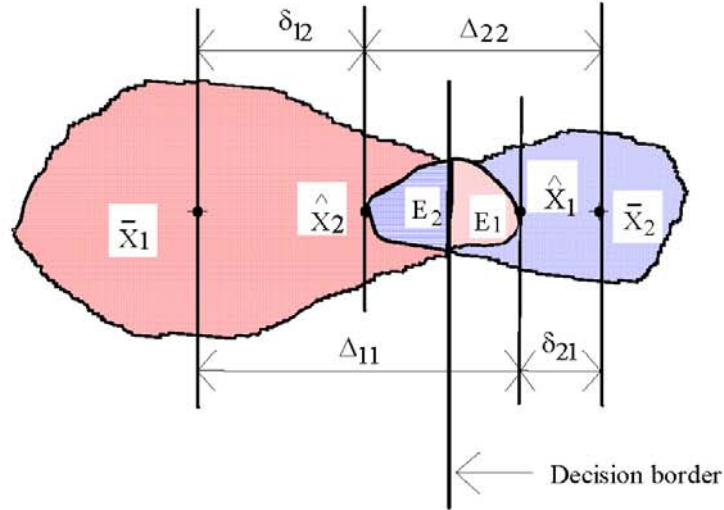


Fig. 2 : Two overlapping clusters

3.2. GENERAL MULTICLASSES CASE

It is important to note that in the previous section, we used two classes to define the classification index function. In the general case where we have more than two classes, we have to reduce the problem to a two-classes problem ω'_1 and ω'_2 after a point is classified so that we can measure the CI function. The nearest class will be taken to give the minimal distance d_{\min} . To define the equivalent second class, we use the mean class of the other classes. Let N be the number of the whole classes. After the point x is classified to a class ω_s , it remains $N-1$ classes. We take the class at which x is classified to be ω'_1 and then we take the rest of the classes (the $N-1$ remaining classes) to compose the second class ω'_2 .

Statistical parameters are then given by the following equations :

$$\bar{x}'_1 = \bar{x}_s \quad \text{and} \quad \Sigma'_1 = \Sigma_s \quad (5)$$

$$\bar{x}'_2 = \frac{1}{N-1} \sum_{j=1, j \neq s}^N \bar{x}_j \quad \text{and} \quad \Sigma'_2 = \frac{1}{N-1} \sum_{j=1, j \neq s}^N \Sigma_j \quad (6)$$

\bar{x}'_1 and \bar{x}'_2 being the mean vectors of the defined two-classes systems and Σ'_1 and Σ'_2 being the corresponding covariance matrices.

Remark: in the general case, regarding to its definition, the CI function can take negative values. In this case, the CI is set to value 0 and the corresponding point is rejected.

Since the two classes ω'_1 and ω'_2 are defined we can now proceed to the classification step by using equation 2, 3 and 4 to fix thresholds and compute CI.

Let r be the number of multiresolution images used in the image segmentation process. Let the reference resolution I_{ref} to the highest (or the lowest) resolution among the r images. Let us order the r images by increasing (or decreasing) resolution according to the choice of the reference image (e.g. $I_1=I_{ref}$, I_2, I_3, \dots, I_r). From the reference resolution I_{ref} through the r images, the process classifies pixels and rejects those which classification is found to be not sufficiently accurate. The accuracy of a pixel classification is obtained by comparing its CI with a pre-defined threshold CI_0 (related to the class at which the pixel is classified). The segmentation procedure is then given by :

- 1) Select r images from the multiresolution system (or pyramid levels) ;
- 2) Set the reference image I_{ref} to the highest (or the lowest) resolution image of the r images;
- 3) For $k = 1$ to $r-1$
 - do
 - Compute the CI's threshold related to each of the classes ;
 - Classify the current pixel P of the current resolution image I_k ;
 - Compute the classification index $CI(P)$ corresponding to this pixel ;
 - If $CI(P)$ is greater than the threshold CI_0^s of the class ω_s at which P is classified to then
 - P 's classification is confirmed
 - else
 - P is rejected waiting for further classification from its "parent" at the next resolution image I_{k+1} .
 - redo
- 4) At the last stage (image I_r), even though pixels have their CI lower than the threshold CI_0^s , their classification is confirmed.

In other words, the segmentation system can be described as follows. The r images having been ordered, start segmentation procedure at image I_1 . During this operation, some pixels may be rejected. Let I'_1 be the image of rejected pixels. At the next step, classify pixels of image I_2 corresponding to I'_1 and reject misclassified ones and let I'_2 be the image containing them. Continue the procedure for the following images until the rejected image I'_k of rank k ($k \leq r$) is empty or until the last image I_r is reached. In this last case, we confirm the classification of every pixel even though its classification index measure is less than the threshold of the class at which it is classified to.

4. EXPERIMENTAL DETAIL AND RESULTS

This segmentation procedure is applied to mosaic of textures to test the performances of ASH algorithm. The test images consist in four different textures as described by fig. 3.a. Textures are taken from Brodatz photographic album⁶ and quantified on 256 grey levels. The different textures are then normalised to have the same statistical properties : same mean value and same variance value. We then build the Gaussian pyramid of the mosaic under study to provide us different resolution images^{1,2,3}. Generally, we used two or three pyramid levels. We have chosen as characterisation features, the fractal parameter H extracted from the fractal Brownian function model^{1,2,3}. We used four features extracted from four different orientations (0° , 45° , 90° and 135°). Figures 3.b to 3.e show results of different steps. Figure 3.b presents an original mosaic of textures (T1=woolen cloth, T2=straw, T3=grass and T4=raffia looped to a high pile). On figure 3.c, we can see a map of CI function corresponding to the original image of figure 3.b. After computing CI, we convert it to an image by quantifying it on 256 grey levels. As it can be seen, the inner region of a texture is uniform but the mean value of the function depends on the texture. It is then a need to define the threshold relatively to each texture. It's also useful to note that the weak values of the CI function correspond to the textures border regions where classification is more difficult. Figure 3.d represents the rejected class (represented in black) after having classified pyramid level 0 (the highest pyramid resolution level). The final segmented image is represented on figure 3.e ; three pyramid levels (0, 1, 2) are used to obtain this result.

This classification method appears to be interesting for two reasons relatively to ASM method. First, images are segmented one after another so the attribute vector is only p long. Secondly, we segment only a fraction of the image from a step to the next one. In a three-levels case, the rejected class I_1 (after having classified level 0) consists of 11.48% of the whole image and I_2 consists of 9.48%. From a step to the next one only a little fraction of the original image has to be computed. This leads to an important computations reduction. More than half of the computation time can then be saved. Anyway it's important to note that the time save depends on the population of the rejected class at each step.

Tables I and II provide us the performances of the two methods in multiresolution textures segmentation. By comparing Table I and table II we can note that ASM keeps best performances but it may be interesting to get ASH performances since one can save important computation time.

r	Levels	T1	T2	T3	T4	Global
	0, 1	99.92	99.91	100.00	100.00	99.96
2	0, 2	99.98	100.00	100.00	99.80	99.94
	1, 2	98.15	100.00	98.48	99.30	98.98
3	0, 1, 2	99.95	100.00	100.00	99.94	99.97

Table I : ASM results for different numbers of levels

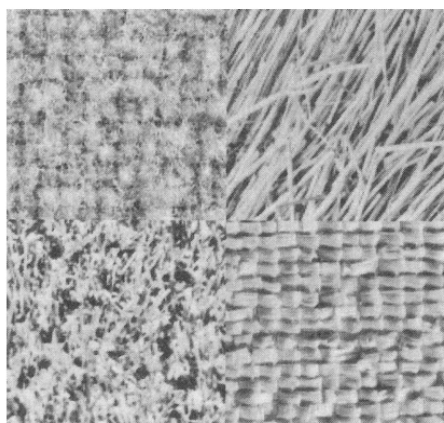
r	Levels	T1	T2	T3	T4	Global
	0, 1	99.98	99.95	94.47	99.98	98.60
2	0, 2	96.96	100.00	88.24	99.94	97.04
	1, 2	93.12	98.48	70.27	99.52	90.35
3	0, 1, 2	99.97	100.00	86.45	99.98	96.60

Table II : ASH results for different numbers of levels

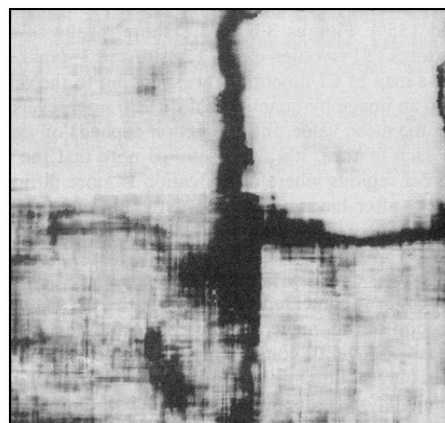
The results presented in these tables correspond to the inner regions of different textures. We exclude then the border effects that influence the results.

T1	T2
T3	T4

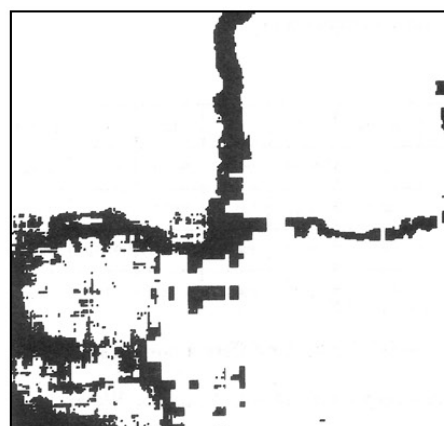
a



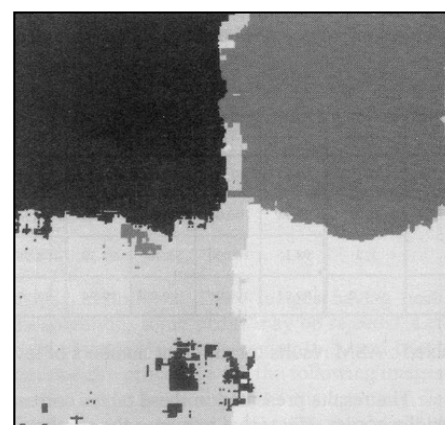
b



c



d



e

Fig. 3 : An original image and the results of different steps in ASH image segmentation.

5. CONCLUSION

We have presented a new method for texture images segmentation in a multiresolution context. This algorithm gives good segmentation results. Even though its performances are inferior than those of ASM, they stay interesting and we can save considerable computation time.

Through several mosaic texture segmentation results, this method is found to be efficient. 90% good segmentation rate is reached. Comparatively to the ASM method, the ASH provides satisfactory segmentation results with a lot of computation time save. Globally, we loose about 5% of segmentation performances but we may win over 50% of computation time.

REFERENCES

- [1] K. Kpalma, V. Haese-Coat and A. Bruno, "Analyse Pyramidale de Textures Naturelles à Travers la Dimension Fractale", 13^{ième} Colloque GRESTI, 16-20 sept. 1991, Juans-les-Pins, France, pp 1053-1056.
- [2] K. Kpalma, A. Bruno and V. Haese-Coat, "Natural Texture Analysis in Multiscale Context Using Fractal Dimension", Proc. SPIE-Visual Commun. and Image Processing, Nov. 1991, Boston, USA, Vol. 1606, pp 55-66
- [3] K. Kpalma, "Analyse Fractale de Textures Naturelles dans un Contexte Multirésolution : Application à la Segmentation d'Images Multirésolution", thèse de Doctorat à l'INSA de Rennes, France, fév. 1992.
- [4] A. G. Weber and A. A. Sawchuk, "Multiple Resolution Image Texture Segmentation", pp 108-120.
- [5] J.C. Simon, "La Reconnaissance des Formes par Algorithmes", Masson, Paris, 1985
- [6] P. Brodatz, "Textures", A photographic Album for Artists and Designers, Dover, New York, 1966.

XIV - Segmentation d'images de télédétection

Véronique Haese-Coat, Kidiyo Kpalma et Joseph Ronsin

Apport comparatif de la texture en segmentation d'images SPOT

Revue Photo-Interprétation,

Editions ESKA

Novembre 1994, pp 49-56 (8 pages)

Résumé

Cet article traite de l'exploitation de l'information de texture pour la classification automatique d'une image SPOT. Le principe des six méthodes de caractérisation de la texture, retenues dans cette étude, est d'abord présenté. Les performances des méthodes en classification sont alors évaluées par référence à la vérité-terrain dont on dispose *a priori*.

F FRANCE**93/4 – 94/1-2
3**

Mission : CNES – SPOT 1 XS2, XS3
 Scène : 032-253 du 04/07/1987, 11h39 TU $\lambda 0,61-0,68 \mu\text{m}$ $\lambda 0,79-0,89 \mu\text{m}$
 Echelle : 1 : 61 000 environ
 Cartes : 1 : 50 000, IGNF feuilles PLOERMEL (X 19) et GUER (XI 19)

APPORT COMPARATIF DE LA TEXTURE EN SEGMENTATION D'IMAGE SPOT APPLICATION À LA FORÊT DE PAIMPONT (I. & V.)

V. HAESE-COAT, K. KPALMA et J. RONSIN

INSA, Laboratoire d'automatique, Équipe Image, partenaire CBI
 20, avenue des Buttes de Coësmes, 35043 Rennes Cedex, France

RÉSUMÉ

Le contenu de cet article traite de l'exploitation de l'information de texture pour la classification automatique d'une image SPOT. Le principe des six méthodes de caractérisation de la texture, retenues dans cette étude, est d'abord présenté. Les performances des méthodes en classification sont alors évaluées par référence à la vérité-terrain dont on dispose a priori.

Nous présentons les résultats d'une étude comparative de différentes méthodes de caractérisation de la texture appliquées à la segmentation automatique d'une image SPOT. L'image retenue pour l'étude est celle de la forêt de Paimpont en mode XS acquise en juillet 1987, au format 512x512 points d'image, dont nous avons extrait le quart SUD-OUEST (taille 256x256), pour lequel une vérité-terrain est disponible. Cette dernière est issue d'une étude réalisée par le COSTEL (réf. article R. BARIOU et al.).

Afin de faciliter la comparaison des résultats entre les différents partenaires du projet, le choix des canaux a été restreint à l'infrarouge moins le rouge, obtenu par différence des canaux proche infrarouge et rouge.

De cette image et de la vérité-terrain associée sont extraites les statistiques correspondantes, c'est-à-dire la répartition des points et des niveaux de gris dans les différentes classes. Ceci fait l'objet du **tableau 1**.

MÉTHODOLOGIE

L'objectif étant la classification automatique de l'image retenue, nous avons scindé la démarche en deux étapes :

- la première correspondant à la phase d'extraction d'attributs caractéristiques de type "texture" à partir de l'image
- la seconde constituant la phase de classification proprement dite, phase au cours de laquelle sont utilisés les attributs précédemment extraits, afin d'atteindre l'objectif initial.

1. MÉTHODES D'EXTRACTION D'ATTRIBUTS

Au vu du contenu texturé de l'image étudiée, les méthodes d'analyse appartenant à la famille des méthodes utilisant les caractéristiques spatiales nous sont apparues comme étant les plus adéquates. Parmi ces dernières, les six méthodes sélectionnées le furent pour les raisons suivantes :

- La méthode des matrices de cooccurrence, parce qu'elle est considérée comme une référence dans le domaine
- La méthode de l'intégrale curviligne,
- La méthode des extrêmes locaux,
- La méthode des histogrammes locaux, toutes trois pour leur simplicité de mise en oeuvre
- La méthode de la décomposition morphologique,
- La méthode de la dimension fractale, ces deux dernières étant choisies pour leur caractère novateur.

Une caractéristique commune à l'ensemble de ces méthodes est qu'elles fournissent un certain nombre d'attributs pour chaque point d'image.

Nous rappelons brièvement ci-après le principe de chacune des méthodes testées.

1.1 Méthode des matrices de cooccurrence [1]

Son principe se fonde sur l'étude des liaisons statistiques des valeurs de niveaux de gris observées en des couples de points. D'où le calcul, pour une relation spatiale donnée du couple de points (distance d entre deux points avec une orientation bien définie), d'un histogramme bidimensionnel du nombre de paires de points de niveaux de gris donnés. A chaque relation spatiale caractérisant les paires de points correspond une matrice de cooccurrence (encore appelée "matrice de dépendance spatiale des niveaux de gris" par certains auteurs), matrice de taille $N \times N$, où N désigne le nombre de niveaux de gris utilisés dans l'image.

De ces matrices, sont extraits des paramètres caractéristiques. Dans notre cas, cette caractérisation s'est limitée aux 6 paramètres les plus couramment utilisés : le *Second moment angulaire*, le *Contraste*, la *Corrélation*, l'*Entropie*, la *Moyenne des lignes* et l'*Ecart-type des lignes*.

Tableau 1 : Statistiques calculées à partir de l'image à étudier et de la vérité-terrain fournie

Classe N°	Type de végétation	Nbre de points	Nbre de parcelles	Nbre moyen de pts/parc.	Moyenne	Écart-type
1	lande boisée	1 706	21	81	150.5	10.9
2	coupe rase	280	4	70	155.3	11.7
3	coupe rase avec début régénération	957	15	64	160.2	12.4
5	conifères régén. naturelle, futaie complète	5 250	71	74	145.1	8.2
7	conifères régén. naturelle, futaie claire	1 160	8	145	142.3	7.2
8	conifères régén. naturelle, futaie hétérogène	1 372	19	72	149.9	8.7
10	bouleaoux-taillis	3 709	47	79	163.3	12.6
11	taillis simple	6 833	56	122	175.7	14.1
12	réserve résineux	2 672	32	84	151.3	11.1
15	réserve feuillus	3 193	24	133	175.1	11.0
16	conifères régén. artificielle, plantations semis fourré	791	9	88	151.6	25.0
17	conifères régén. artificielle, plantations semis fourré	2 043	24	85	170.5	14.8
18	conifères régén. artificielle, plantations semis fourré	1 590	14	114	165.0	8.1
19	conifères régén. artificielle, plantations semis fourré	332	3	111	154.1	5.8
20	conifères régén. naturelle, plantation semis fourré+gaulis	147	1	147	146.4	5.0
21	conifères régén. naturelle, bas et haut perchis	271	7	39	148.4	6.3
22	feuillus, plantations semis fourré	22	1	22	161.7	7.5
23	feuillus et conifères, plantations semis fourré	31	1	31	140.2	19.9
24	feuillus, haut perchis	489	8	61	175.9	12.4
25	feuillus, futaie	648	9	72	176.9	7.2

1.2 Méthode de l'intégrale curviligne [2]

Le principe de la méthode repose sur l'intégration curviligne du signal de radiométrie le long de la ligne de balayage L_i centrée sur le pixel à traiter et ayant une longueur et une orientation θ_i données.

Soient (x,y) les coordonnées du pixel traité dans le plan de l'image, $z = f(x,y)$ le niveau de gris de ce point. En approximant la variation du niveau de gris entre deux pixels par un segment de droite et en utilisant une métrique euclidienne grâce à l'introduction d'un facteur d'échelle entre la variation spatiale (dx,dy) et la variation de niveau de gris dz , le calcul de l'intégrale curviligne le long de la ligne L_i revient à :

$$\int_{L_i} ds \quad \text{avec } ds = (\lambda^2 (dx^2 + dy^2) + dz^2)^{1/2}$$

Le calcul de l'intégrale curviligne le long de L_i s'arrête lorsque la valeur de l'intégrale atteint la valeur μ du seuil fixé. Le paramètre caractéristique est alors la longueur parcourue le long de L_i exprimée en nombre de points.

Une autre façon de procéder consiste à fixer la distance à parcourir le long de la droite L_i . Le paramètre caractéristique alors obtenu est la valeur de l'intégrale calculée le long de L_i (longueur du signal de luminance pour un support donné).

Le nombre de paramètres caractéristiques extraits en chaque point par cette méthode dépend directement du nombre d'orientations θ_i balayées. Ce nombre est généralement fixé égal à 4 ($\theta = 0, 45, 90$ et 135°).

1.3 Méthode des extréma locaux [3]

Son principe, pour une ligne de balayage L_i et dans une direction θ_i donnée, repose sur l'évaluation de la fréquence relative des extréma locaux dans les niveaux de gris, c'est-à-dire la fréquence relative des pics de radiométrie, comme mesure principale pour caractériser les images.

L'algorithme développé comporte deux étapes :

- ♦ La première correspond à un processus de lissage destiné à éliminer les pics de radiométrie de faible amplitude afin de ne garder que les extréma principaux.

- ♦ La deuxième phase correspond à la détection et au comptage des extréma principaux.

De nouveau, le nombre de paramètres extraits découle directement du nombre d'orientations θ_i retenues. Pour cette étude, il est fixé égal à 4.

1.4 Méthode des histogrammes locaux [4]

Cette méthode consiste à extraire l'information caractéristique de l'image à partir des histogrammes locaux de niveaux de gris. Chaque histogramme obtenu sur une fenêtre de taille $n \times n$ ($= n^2$ points d'image ; n étant ici choisi égal à 15) est défini comme un vecteur caractérisé par sa norme (ou module) et sa phase définies de la manière suivante. Les niveaux de gris étant quantifiés sur r niveaux, alors :

- le module d'un histogramme local est défini comme la distance entre l'histogramme particulier $h_0 = \{n^2/r, n^2/r, \dots, n^2/r\}$, et l'histogramme local h considéré.

$$\|h\| = D(h_0, h) = \sum_{i=1}^r \left(r_i - \frac{n^2}{r} \right) \log(r_i) \geq 0$$

- quant à la phase, c'est l'indice j le plus petit de l'histogramme maximal $h_j^M = \{0, 0, \dots, n^2, \dots, 0\}$ le plus proche au sens de D de l'histogramme local considéré. Autrement dit, j est l'indice du maximum de l'histogramme local :

$$\Phi(h) = j \quad \left(\text{en unités } \frac{2\pi}{r} \right)$$

Chaque point d'image est donc caractérisé par deux paramètres : le module et la phase de l'histogramme local calculés au voisinage du point.

1.5 Méthode de la décomposition morphologique [5]

La morphologie mathématique fut initialement étudiée par G. Matheron et J. Serra. L'idée de base de la morphologie mathématique est de comparer l'objet à analyser à un objet de forme connue appelé *élément structurant* [8, 9]. Les opérations de base sont l'*érosion*, la *dilatation*, l'*ouverture* et la *fermeture*. L'utilisation de ces opérations nécessite le choix au préalable d'un élément structurant dont on étudie l'impact sur l'objet d'intérêt. La méthode de décomposition présentée, ici, utilise l'opération d'ouverture qui est l'enchaînement d'une érosion suivie d'une dilatation. Pour des images multi-niveaux (c'est le cas ici), une érosion consiste à prendre la valeur minimale des points en coïncidence avec l'élément structurant ; la dilatation, quant à elle, revient à prendre la valeur maximale parmi ces mêmes points.

Dans notre méthode, le principe repose sur l'utilisation d'une famille d'éléments structurants de différentes tailles, possédant une forme carrée ($B_0 = \text{carré } 1 \times 1$, $B_1 = \text{carré } 3 \times 3$, $B_2 = \text{carré } 5 \times 5$, ...).

L'algorithme de la décomposition est alors le suivant : l'ouverture de l'image $f(x,y)$ par l'élément structurant B_n , le plus grand dans le groupe est d'abord calculée. Toutes les primitives plus larges que cet élément structurant sont conservées dans l'image ouverte désignée par $S_0(x,y)$, et toutes les primitives de taille inférieure se voient complètement éliminées. Les primitives un peu moins larges que B_n peuvent alors être extraites par l'ouverture de l'image différentielle $f(x,y) - S_0(x,y)$, réalisée cette fois avec le second plus grand élément structurant B_{n-1} . Ensuite et de la même façon, les primitives de moins en moins larges sont extraites en utilisant des éléments structurants de plus en plus petits.

A l'issue de cette décomposition, des paramètres caractéristiques sont calculés sur chaque image composante et forment les attributs de la texture. On peut citer, comme paramètres extraits, les moyennes de niveaux de gris calculées sur des fenêtres locales qui, bien que simples, donnent de bons résultats en segmentation.

Le nombre de paramètres extraits par cette méthode dépend de la profondeur de la décomposition. Dans nos expériences, l'image initiale a été ouverte en utilisant successivement les éléments structurants carrés de taille 13×13 , 9×9 , 5×5 et 1×1 , soit 4 paramètres extraits pour chaque point d'image.

1.6 Méthode de la dimension fractale [6]

Nous ne rappellerons pas ici la théorie des fractales développée par Mandelbrot [7]. Disons simplement que les mouvements browniens fBm (fractional Brownian motion) ont été introduits par Mandelbrot pour décrire les processus stochastiques de type gaussien. Le fBm est essentiellement caractérisé par un paramètre H qui définit sa dimension fractale. Si D_F et D_T sont les dimensions fractale et topologique d'un fBm paramétré par la valeur H , on a alors :

$$D_F = D_T + 1 - H$$

Puisque H est défini sur le domaine $]0,1[$, il est clair que la dimension fractale D_F est fractionnaire et strictement supérieure à la dimension topologique D_T .

En pratique, le calcul de H est basé sur la moyenne des carrés des écarts des niveaux de gris entre les pixels situés dans une fenêtre d'analyse centrée sur le pixel x et ceux situés dans une autre fenêtre de même taille, centrée sur le pixel $x + \Delta x$, donc distante de $\|\Delta x\|$ de la première. En prélevant une série de valeurs de l'espérance mathématique $E[(I(x + \Delta x) - I(x))^2]$ pour différentes valeurs du vecteur déplacement Δx , l'estimation de la droite passant par les points ainsi obtenus, dans un référentiel log-log, donne le paramètre H qui représente la pente de cette droite. La valeur de la pente ainsi mesurée est alors attribuée au pixel x .

Cette méthode a été modifiée pour prendre en compte le caractère anisotrope du signal d'image. On estime donc le paramètre H pour plusieurs déplacements Δx et suivant différentes orientations.

Le nombre de paramètres caractéristiques dépend ici encore directement du nombre d'orientations choisies. Compte tenu de la lourdeur de cette méthode du point de vue coût calcul, nous nous sommes limités aux quatre orientations principales ($0, 45, 90$ et 135°).

2. CLASSIFICATION

Les paramètres de texture ainsi obtenus sont ensuite soumis à un classifieur bayésien qui effectue la classification de l'image. Cette étape se décompose en réalité en deux phases bien distinctes :

- l'apprentissage sur des zones précises de l'image, représentatives de chaque classe présente dans l'image,
- la classification proprement dite, qui utilise les résultats de l'apprentissage pour classifier l'image point par point.

RÉSULTATS

L'expérience présentée ici concerne la classification des 20 classes recensées dans l'image originale.

Un certain nombre d'essais préalables nous ont conduit à définir une procédure de classification mieux adaptée au

problème posé par une vérité-terrain comportant des parcelles très petites (quelques dizaines de points par parcelle, comme le montre le **tableau 1**). Les modifications apportées ont pour but de s'affranchir des problèmes liés aux frontières entre parcelles. Elles se situent à trois niveaux différents :

1. LORS DE LA PHASE D'APPRENTISSAGE

A ce niveau, la modification consiste à n'utiliser comme échantillons (points d'image) pour effectuer l'apprentissage que ceux situés en zones homogènes, c'est-à-dire éloignés des zones de frontières entre parcelles. Pour ce faire, on dresse une carte des points autorisés. Un point appartient à cette carte si, dans une fenêtre de taille 5 X 5 centrée sur ce point, n'existe aucun point frontière. Pour un point du bord de l'image, on ne tient compte que de la portion de fenêtre de taille 5 X 5 s'inscrivant dans l'image. Le choix de la taille de fenêtre résulte d'un compromis entre degré d'homogénéité de la zone et nombre de points conservés. Le processus se déroule alors de la façon suivante : on tire un échantillon pour la classe i ; s'il est présent dans la carte des points autorisés, on le conserve ; sinon on effectue un nouveau tirage.

2. LORS DE LA PHASE DE CLASSIFICATION

Le point n'est classifié que s'il appartient à la classe des points autorisés.

3. LORS DE LA PHASE DE VISUALISATION DES RÉSULTATS DE CLASSIFICATION

Pour faciliter l'interprétation de ces résultats, on introduit un post-traitement consistant, pour chaque parcelle repérée au cadastre, à lui affecter le numéro de la classe majoritaire dans la parcelle à l'issue de la phase de classification.

Nous avons expérimenté toutes les méthodes de caractérisation retenues pour l'étude, en y ajoutant l'utilisation des seuls paramètres de radiométrie des canaux proche infrarouge moins rouge et rouge seul.

La classification s'est limitée à 18 classes au lieu des 20 recensées, l'apprentissage ne pouvant se faire sur les classes n° 22 et n° 23 du fait du trop petit nombre de points leur appartenant (22 points dans la classe n° 22 et 31 points dans la classe n° 23 comme l'indique le **tableau 1**).

Les **tableaux 2** et **3** ci-après fournissent les résultats complets de cette expérience et ce, par méthode testée. Le **tableau 2** donne les résultats obtenus pour chaque classe ; ces résultats sont des pourcentages de points bien classifiés par comparaison avec la vérité terrain fournie. Certains chiffres sont semblables pour différentes méthodes ; ceci est tout à fait normal compte tenu du post-traitement effectué sur les résultats bruts de classification et qui conduit à un classement par "parcelles" plutôt qu'à un classement "point par point".

Au regard du **tableau 2**, on constate que certaines classes sont peu ou jamais détectées ; c'est le cas des classes :

- 1, 3 (improductifs : lande boisée, coupe rase avec début de régénération)
- 10, 11, 15 (taillis)
- 5, 12, 18, 21 (divers).

D'autres, au contraire, sont bien détectées quelle que soit la méthode utilisée. Il s'agit des classes :

- 2 (terre à nu, coupe rase)
- 7, 8, 16, 17, 19, 20 (conifères régénération naturelle ou artificielle)
- 24, 25 (feuillus).

Les autres classes sont parfois détectées par l'une ou l'autre des méthodes mais aucune règle ne semble s'affirmer.

L'analyse des résultats globaux consignés dans le **tableau 3** montre qu'en définitive, même si certaines classes obtiennent des pourcentages de bonne classification voisins de 100%, elles ne représentent qu'un faible nombre de points de l'image-test.

Le **tableau 3**, quant à lui, nous fournit des résultats globaux, c'est-à-dire toutes classes confondues.

Que l'on exprime les résultats en nombre de points bien classés ou en nombre de parcelles bien classées, le pourcentage résultant ne dépasse pas les 25%, soit un quart de l'image bien classifié dans le meilleur des cas : ce qui est peu. Ce tableau permet également de comparer directement les performances des six méthodes retenues. La méthode de la décomposition morphologique apparaît comme la plus performante alors que celles de l'intégrale curviligne et des fractales sont les plus décevantes. Une explication peut résider dans le fait que les supports de caractérisation employés ici étaient petits, compte tenu de la faible taille des parcelles à caractériser. D'où l'obtention de caractérisations insuffisamment pertinentes vis-à-vis d'un processus de classification.

On notera néanmoins que l'utilisation des seuls paramètres de radiométrie (colonne lum.) conduit à de moindres performances que celles obtenues par les meilleures méthodes de caractérisation de textures.

La **figure 1** visualise les résultats de classification obtenus dans le cadre de l'utilisation de la méthode des matrices de cooccurrence (b), de la méthode de la décomposition morphologique (c) et des extréma locaux (d).

CONCLUSION

Parmi les six méthodes de caractérisation de texture retenues, celles utilisant les caractéristiques de radiométrie (*méthodes de la décomposition morphologique, des matrices de cooccurrence et des extréma locaux*) donnent de meilleurs résultats que celles travaillant sur les mesures de variations de radiométrie (*méthodes de l'intégrale curviligne ou des fractales*). Ceci renforce l'idée que l'image testée n'a finalement pas un contenu texturé très évident. L'œil humain d'ailleurs n'y distingue guère plus de 5 classes de végétation différentes.

Une approche purement "texture" est donc sans doute trop ambitieuse vu la nature de l'image à tester. La caractérisation de la texture se heurte à la trop grande finesse du parcellaire. La restriction à l'analyse de l'image de la différence des canaux proche infrarouge et rouge appauvrit l'information. Enfin, la résolution des images SPOT fournies est probablement trop grossière pour que l'on puisse fonder une classification uniquement sur l'approche texture.

Une meilleure solution résiderait dans une phase préalable de segmentation s'appuyant sur les contours de luminance présents dans l'image. Une fois l'image segmentée en régions homogènes, la caractérisation de la texture de chacune de ces régions apporterait alors l'information supplémentaire indispensable à la réussite de la phase de classification.

Il est bien évident que la validation de cette approche sur un éventail plus large d'images satellitaires s'impose.

Tableau 2 : Résultats de classification des 18 classes originales selon la méthode de caractérisation utilisée

% de bonne classification							
Classe n°	Cooc.	Integ.	Morpho.	Fractale	Extrêma	Histog.	Lum.
1	0.	0.	0.	0.	0.	0.	0.
2	0.	67.0	67.0	67.0	100.0	67.0	
3	0.	0.	0.	0.	0.	0.	0.
5	26.6	0.	4.5	0.	0.	0.	0.
7	76.8	0.	76.8	89.2	0.	93.4	93.4
8	31.8	15.8	83.5	0.	48.3	0.	67.4
10	21.6	0.	60.9	0.	0.	0.	3.5
11	14.0	12.8	6.8	0.	77.3	1.5	12.3
12	0.	0.	0.	0.	0.	0.	0.
15	24.4	0.	34.2	0.	0.	0.	29.2
16	50.5	50.5	77.6	100.0	0.	50.5	50.5
17	54.1	0.	25.7	0.	0.	40.8	40.8
18	0.	22.7	29.1	0.	9.3	0.	0.
19	0.	0.	100.0	0.	100.0	100.0	0.
20	100.0	0.	100.0	0.	0.	100.0	100.0
21	0.	0.	0.	0.	0.	39.1	0.
24	59.3	0.	59.3	0.	0.	59.3	59.3
25	83.4	83.4	83.4	83.4	83.4	83.4	100.0

Tableau 3 : Résultats globaux de classification

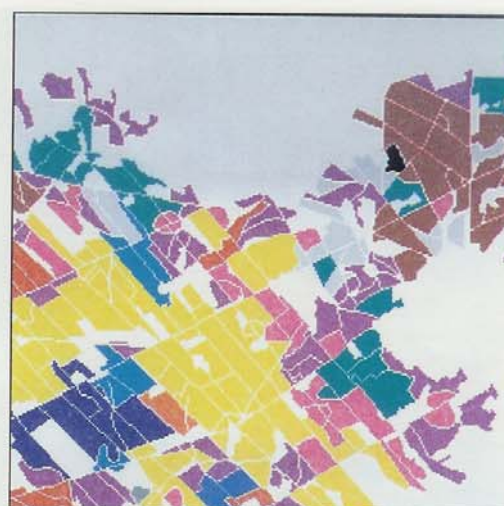
MÉTHODES							
Résultats	COOC.	INTEG.	MORPHO	FRACTALE	EXTREMA	HISTOG.	LUM.
nbre de points à classer	33 443						
nbre de points bien classés	6 988	2 324	8 185	2 362	6 655	3 788	5 809
%	20.9	6.9	24.5	7.1	19.9	11.3	17.4
nbre de parcelles à classer	273						
parcelles bien classées	61	22	68	18	46	30	45
%	22.3	8.1	24.9	6.6	16.8	11.0	16.5

RÉFÉRENCES

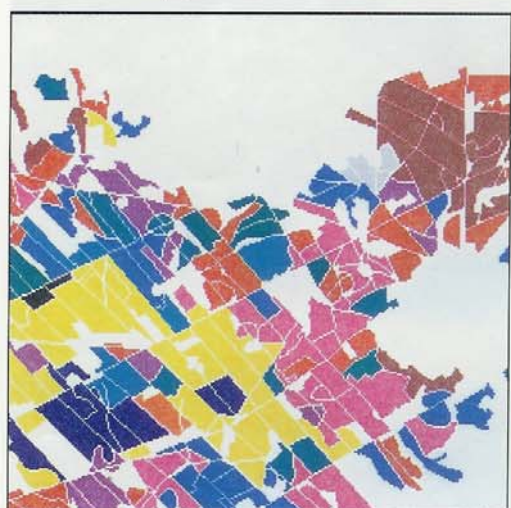
- [1] HARALICK R. et AL., 1973, "Texture feature for image classification", *IEEE Trans. Syst., Man and Cybernetics*, Vol. SMC-3, pp. 610-621.
- [2] BARBA D. et RONSIN J., Sept. 1983, "New method in texture analysis in the context of image segmentation", *Proc. 2nd European Signal Processing Conference*, West Germany, pp. 283-286.
- [3] MITCHELL O.R. et CARLTON S.G., 1978, "Image segmentation using a local extrema texture measure", *Pattern Recognition*, Vol. 10, pp. 205-210.
- [4] LOWITZ G.E., 1983, "Can a local histogram really map texture information", *Pattern Recognition*, Vol. 16, n° 2, pp. 141-147.
- [5] WANG D., Déc. 1991, "Décomposition morphologique multi-tailles et segmentation adaptative de textures", *Thèse de Doctorat*, INSA de Rennes.
- [6] KPALMA K., Fév. 1992, "Analyse fractale de textures naturelles dans un contexte multirésolution : Application à la segmentation d'images multirésolution", *Thèse de Doctorat*, INSA de Rennes.
- [7] MANDELBROT B.B., 1982, "The fractal geometry of nature", San Francisco, CA : Freeman.
- [8] SERRA J., 1982, "Image analysis and mathematical morphology", Academic Press, New York.
- [9] MATHERON G., 1975, "Random sets and integral geometry", J. Wiley, New York.



(a) vérité-terrain
(a) terrain truth
(a) realidad de terreno



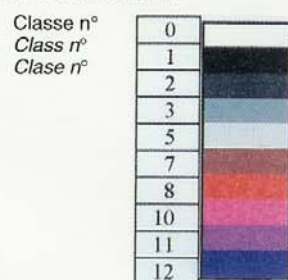
(b) cooccurrence
(b) co-occurrence
(b) concomitancia



(c) décomposition morphologique
(c) morphological breakdown
(c) descomposición morfológica



(d) extréma locaux
(d) local extrema
(d) extremos locales



(e) palette utilisée
(e) colour range used
(e) paleta utilizada

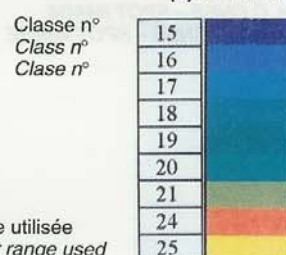


Figure 1 : Résultats de classification
Figure 1 : Classification results
Figura 1 : Resultados de clasificación

Voir légende des couleurs dans le tableau 1, page 50.
See colour key in table 1, page 56.
Ver índice de colores en la tabla 1, pagina 61.

F FRANCE	CNES SPOT 1 XS scène 032-253	04/07/1987 11H39 TU	1 : 61 000 environ $\lambda 0,61-0,68 \mu m$ $\lambda 0,79-0,89 \mu m$	1 : 50 000, IGNF feuilles PLOERMEL (X 19) et GUER (XI 19)
-----------------	---------------------------------	------------------------	--	---

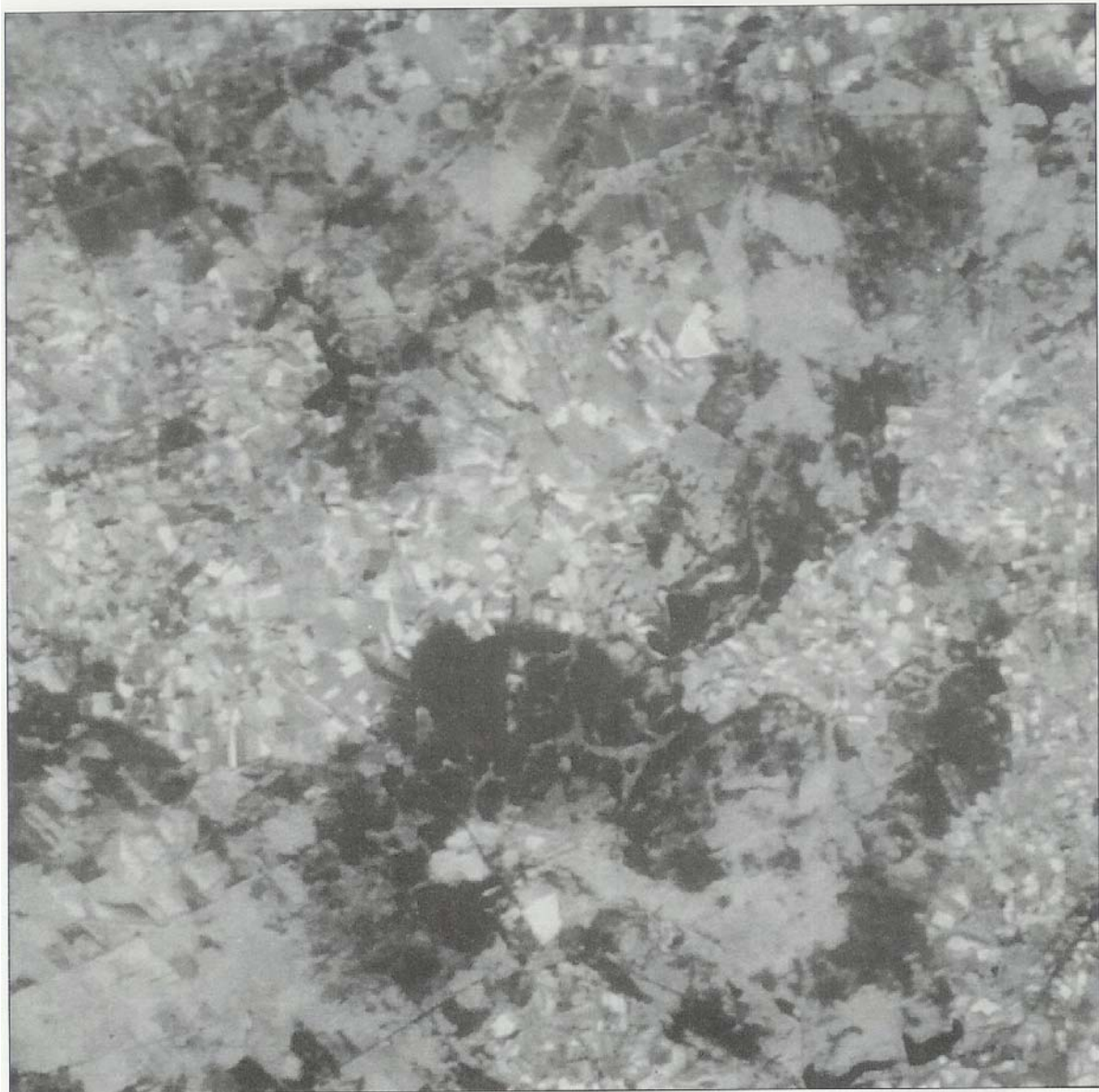


Figure 4 : Forêt de Paimpont, canal XS3
Figure 4 : Paimpont Forest, C CNES – SPOT IMAGE
Figura 4 : Bosque de Paimpont, C CNES – SPOT IMAGE

XV - Recalage d'images multidate et multicapteur

Olivier Thépaut, Kidiyo Kpalma and Joseph Ronsin

Automatic registration of ERS and SPOT multisensor images in a data fusion context,

Elsevier's Forest Ecology and Management,

January 2000

Vol. 128, pp 93-100 (8 pages)

Résumé

Cet article présente une méthode de recalage automatique d'images multidate et multicapteur de télédétection comme prétraitement dans un processus de segmentation consensuelle d'images multicanal. Cette méthode à deux étapes consiste en la correction géométrique suivie par la compensation de la translation résiduelle réalisée en détectant le maximum de l'intercorrélation entre les images multidate. Cette deuxième étape est précédée par une détection de contours appliquée aux images préfiltrées du satellite européen ERS (European Remote sensing Satellite) et du satellite multicapteur SPOT (Satellite Pour l'Observation de la Terre). Le recalage de ces deux types d'images constitue, en effet, l'incontournable étape initiale dans un schéma de fusion de données multicanal qui exige un recalage d'images par une mise en correspondance point à point.



ELSEVIER

Forest Ecology and Management 128 (2000) 93–100

 Forest Ecology
and
Management

www.elsevier.com/locate/foreco

Automatic registration of ERS and SPOT multisensor images in a data fusion context

Olivier Thépaut*, Kidiyo Kpalma, Joseph Ronsin

INSA, Laboratoire ARTIST, 20, avenue des Buttes de Coësmes, CS 14315, 35043 Rennes Cedex, France

Accepted 24 September 1999

Abstract

This paper deals with an automatic registration method of multirate and multisensor remote sensing images as preprocessing in a multichannel consensual segmentation scheme. This two step method consists in a geometrical correction followed by residual translation compensation, performed by detecting the crosscorrelation maximum between multirate images. This second step is preceded by an edge detection for prefiltered European remote sensing satellite (ERS) and satellite pour l'observation de la terre (SPOT) multisensor images. The registration of these two kinds of images constitutes, indeed, the unavoidable first part of a multichannel data fusion scheme that requires point to point matching. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Environment management; Data fusion; Multisensor image registration

1. Introduction

Faced with the diversity of remote sensing systems and of the data they produce, there is a need to synthesize this information abundance. ERS SAR (synthetic aperture radar) instruments actively generate images in the field of microwaves. As for SPOT HRV (high-resolution visible) instruments, they passively generate images in the visible and near infrared field. To those two remote sensing systems correspond two kinds of images containing redundant and complementary data. Since wavelengths differ, acquired information differs too. While SPOT images give high

precision information about agricultural and forest patchwork, road and fluvial network locations, ERS images give information about soil and vegetation nature and moisture and, moreover, their acquisition is much less dependent on weather conditions or sunlight.

Unfortunately, these images do not match directly. The acquisition geometry features infer a particular processing for each kind of image to achieve multichannel registration. The drawback of registration methods based on the detection followed by the matching of points or structures of control is that, they require a more or less automatic matching of the control entities (Rignot et al., 1991; Brown, 1992). And those developed for images acquired by the same kind of sensor cannot be utilized for heterogeneous ones (Djardji et al., 1993).

* Corresponding author. Tel.: +33-299-286-650; fax: +33-299-286-495.
E-mail address: othepaut@insa-rennes.fr (O. Thépaut).

The purpose of this paper is to present our multidade and multisensor registration method consisting in a geometrical correction followed by a spatial cross-correlation (Thépaut et al., 1998). Its principal advantage is its reliability and its full automation. After their registration, superposable scenes can be used to extract consensual segmentation, performed by a cooperative segmentation step followed by a decision one.

We also present here the qualitative and quantitative results of our automatic registration method and an immediate application of speckle cooperative multi-channel filtering that necessitates coregistered images.

2. Geometrical correction

Geometrical correction aims at replacing the images in a common reference system. This will constitute the first step of our multidade and multi-sensor registration method. We assume in this section that the acquired scenes correspond to a plate area at sea level and, then, do not require the use of a digital elevation map (DEM) to get rid of foreshortening, layover and shadowing.

2.1. Acquisition model

Depending on the sensor, SAR for ERS and HRV for SPOT, the acquisition geometry is not the same (Curlander and McDonough, 1991; Bonn and Rochon, 1992). Some of its variable parameters are shown in Fig. 1: the satellite altitude H , its distance from the target R , the incidence, look and site angles, α , θ , and ψ , respectively, and the reference ellipsoid. Consequently, the resulting images are not immediately superposable.

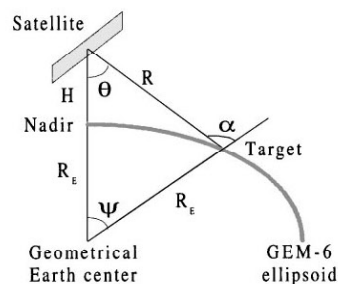


Fig. 1. Acquisition geometry.

The Goddard earth model 6 (GEM-6) ellipsoid is defined by the relation between the earth radius R_E , the geodesic latitude φ , and the equatorial and polar earth radii a and b , respectively

$$R_E = a \sqrt{\frac{\cos^2 \varphi + (b/a)^4 \sin^2 \varphi}{\cos^2 \varphi + (b/a)^2 \sin^2 \varphi}}$$

Then, the distance nadir-target being small enough, compared to the earth radius, to consider that the earth radii at the nadir and at the target are identical, the satellite altitude H , the look angle θ and ψ , are given by

$$H = \sqrt{R_E^2 + R^2 + 2R_ER \cos \alpha} - R_E,$$

$$\sin \theta = \frac{R_E}{R_E + H} \sin \alpha, \quad \psi = \alpha - \theta$$

The image surface is assumed to be locally represented by an ellipsoid. This excludes any local surface slope, which can be important for hilly regions. Then, from the geometrical model and the data accompanying them, it is possible to resample the images to a common reference system.

2.2. Geometrical correction

Many preprocessing levels are available, nevertheless, in a context of data fusion, the less preprocessing there is the better. This is why the ERS SAR PRI (precision image) and SPOT P-1A (Panchromatic 1A) images have been chosen.

These two products require their own processing that will be presented in this section, and of which the principle is illustrated in Fig. 2.

ERS SAR PRI images are in ground range, meaning the earth's roundness has already been compensated.

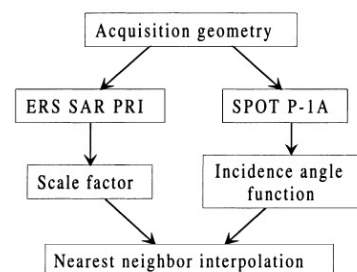


Fig. 2. Geometrical correction scheme.

A simple scale factor and a nearest-neighbor interpolation are sufficient to ensure the registration to a common geodesic grid in latitude and longitude.

As for SPOT P-1A images, they require a processing, according to satellite-target distance, able to compensate the distortion resulting of the earth geometry. Modelling the acquisition reference system, simple trigonometric relations ensure the commutation to the chosen reference system. After this processing, a simple nearest-neighbor interpolation is sufficient to perform the resampling to fit the ERS image resolution (12.5 m).

3. Registration

Nevertheless, some residual translations between the images resulting from geometrical correction processing usually remain and need to be compensated.

3.1. Multidate registration

A first registration scheme concerns images acquired by the same sensor at different dates, as presented in the darkest part of Fig. 3. Since the acquisition method is identical, the multidate images contain the same kind of data. Thus, a simple pixel-by-pixel spatial crosscorrelation ρ of a $m \times m$ sub-image f on a $n \times n$ sub-image g ($n > m$) enables the detection

of a correlation peak ρ_{\max} , that gives the value of the residual translation between geometrically corrected images

$$\rho(u, v) = \frac{\sum_{x=0}^{m-1} \sum_{y=0}^{m-1} f(x, y) g(u+x, v+y)}{\sqrt{\sum_{x=0}^{m-1} \sum_{y=0}^{m-1} f^2(x, y) \sum_{x=0}^{m-1} \sum_{y=0}^{m-1} g^2(x, y)}}$$

3.2. Multisensor registration

The multirate registration principle cannot be directly extended to multisensor registration. The ERS SAR and SPOT HRV sensors do not produce the same kind of data, and therefore, a simple correlation peak seeking is nonsensical.

The heterogeneity of multisensor data, that constitutes the high interest of combining them, induces also a high difficulty in matching them. While the passive SPOT HRV acquisition simply generates images corrupted by low additive noise, the active ERS SAR acquisition generates images which are highly corrupted by multiplicative noise, speckle, result of the constructive and destructive interference of the elementary targets in the resolution cells. For an intensity image, the multiplicative model is defined by Eq. (1), where I , r and n represent, respectively, the intensity, the noise-free reflectivity and the speckle itself (Baraldi and Parmiggiani, 1995)

$$I = rn \quad (1)$$

Let note L the equivalent number of looks of an ERS SAR intensity image. It is calculated from the Eq. (2) in a homogeneous zone, where the coefficient of variation C_r of the reflectivity r is null and speckle is partially developed (by opposition to the fully developed speckle in single-look images)

$$C^2 = \frac{\sigma^2}{\mu^2} = C_n^2 = \frac{\sigma_n^2}{\mu_n^2} = \frac{1}{L} \quad (2)$$

Where μ_n is the mean, σ_n the standard deviation and C_n the coefficient of variation of the speckle n , and μ , σ and C the ones of the intensity I , which probability density function (PDF) model is a gamma distribution in a homogeneous zone

$$p(I) = \frac{L^L}{2^L \sigma^2 L \Gamma(L)} \exp\left(-\frac{LI}{2\sigma^2}\right) I^{L-1} \quad (3)$$

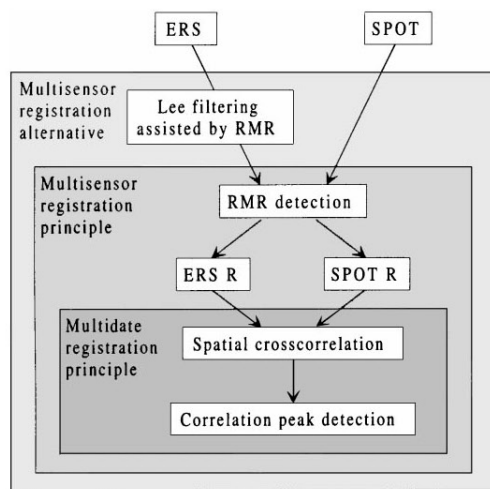


Fig. 3. Multisensor registration alternative steps.

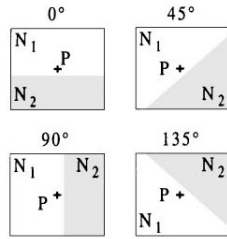


Fig. 4. Edge detector orientations.

3.2.1. Principle

A solution consists in seeking the correlation peak using edge information rather than radiometry data. Though some of the edges detected in ERS images stem from a strong variation of soil moisture or vegetation growth, we assume that a scene contains a higher number of stable than evanescent structures. The mid-gray part of Fig. 3 illustrates this principle whose first advantage resides in its simplicity.

To extract the edge data, we have chosen an adaptive edge detector based on the computation of the radiometric mean ratio (RMR) on windows which are adaptive in size (from 3×3 to 9×9 pixels) and in edge orientation (0° , 45° , 90° and 135°) as illustrated in Fig. 4 (Touzi et al., 1988).

From the smallest to the largest size of window centered at p point and for the four directions, the RMR of the sub-windows, containing N_1 and N_2 pixels, is computed. Its minimal value for a given window size is compared to the probability of false alarm (PFA) of an edge detection in a homogeneous zone. If it is inferior, the process is stopped, else, it is continued with a larger window. While the smaller windows detect thin details, the larger ones are better adapted to detect the coarser details.

One of the major advantages of the RMR detector is that the detection is made at constant PFA in ERS SAR images. The PFA correspond to the probability of detecting an edge in a homogeneous zone. Let note R is the ratio of the intensity means μ_1 and μ_2 on sub-windows N_1 and N_2

$$\begin{cases} R = \mu_2/\mu_1 & \text{if } \mu_1 \geq \mu_2 \\ R = \mu_1/\mu_2 & \text{if } \mu_1 < \mu_2 \end{cases}$$

Then, the PDF of R on a homogeneous zone is deduced from Eq. (3)

$$p(R) = \frac{\Gamma(N_1L + N_2L)}{\Gamma(N_1L)\Gamma(N_2L)} \times \left\{ \left(\frac{N_2}{N_1} \right)^{N_2L} \frac{R^{N_1L-1}}{(R + (N_2/N_1))^{(N_1+N_2)L}} + \left(\frac{N_1}{N_2} \right)^{N_1L} \frac{R^{N_2L-1}}{(R + (N_1/N_2))^{(N_1+N_2)L}} \right\}$$

For a given decision, threshold T , the PFA of R is

$$P_1(T) = \int_0^T p(R) dR$$

And the PFA of R for the four directions is

$$P_4(T) = 1 - (1 - P_1(T))^3$$

This relation allows the determination of the theoretical detection thresholds T , to obtain the chosen PFA.

This RMR edge detector is also applied for HRV SPOT images, considering the following assumption. SPOT images can be seen as ideally filtered ERS images.

When L is about 3 in ERS SAR PRI images, we will suppose that its value is $\lambda \gg 3$ in SPOT HRV P-1A images. Then, for a given PFA, the thresholds T will be very close to 1. However, for a high value of L and then for N_1L or N_2L , the gamma function is hardly computable. Therefore, we use the same thresholds T for ERS and SPOT images as much to simplify the problem than the algorithm. Thus, the use of the same T values induces a much more severe processing for SPOT than for ERS images.

It can also be noted that we obtain multilevel edges, giving a fuzzy location of the real edges.

The edge data do not depend on the sensor; consequently, the subsequent processing is the same as in multirate. Unfortunately, this method is simple but not entirely reliable, because it produces detection errors due to the high noise corruption of ERS images, so we propose an alternative.

3.2.2. Alternative

It consists in lowering the PFA in the ERS image homogeneous zones by applying an adaptive Lee filter, as presented in the lighter part of Fig. 3. The principle of the adaptive edge detector is extended to the detection of lines and isolated backscatters (Lopès et al., 1993).

Instead of filtering with a gamma gamma maximum a posteriori (GGMAP) filter, the homogeneous areas detected at a constant PFA, a Lee filter is used which is less time costly, and for which the estimated unspeckled radiometry is given by Eq. (4) (Lee, 1980; Baraldi and Parmiggiani, 1995)

$$\hat{r} = \frac{I}{\mu_n} \frac{C^2 - C_n^2}{C^2 + C_n^2} + \frac{\mu}{\mu_n} \frac{C_n^2 + C_n^4}{C^2 + C_n^4} \quad (4)$$

Since in the case of a 3-look intensity image in a homogeneous zone, $\mu = 1$ and $\sigma = 1/L$, Eq. (4) can be written

$$\hat{r} = I \left(1 - \frac{C_n^2}{C^2} \right) + \mu \frac{C_n^2}{C^2}$$

While the homogeneous zones are filtered, the edges, lines and isolated backscatters are preserved. It thus, follows that the RMR detector, applied to the adaptive Lee filtered images, provides more reliable information on edge location, because the homogeneous zones are less noisy than in the original images. Multisensor processing is then carried out in the same way as the multirate one.

The registration step is unavoidable, therefore, our simple and reliable multirate and multisensor registration method constitutes a robust automatic preliminary step of our data fusion scheme. The pixel to pixel matching of the resulting images allows us to consider data fusion schemes such as consensual segmentation, classification, temporal agriculture and forest supervision or soil hygrometry fusion.

4. Results

To illustrate our purpose, we present some results of our two step registration method and an application example.

4.1. Geometrical correction

Fig. 5 presents two examples of geometrical correction of ERS and SPOT images on a common reference system in latitude and longitude. It concerns the polder zone, neighboring the Mont-Saint-Michel in France, that is visible in the upper right corner. This plate area at sea level does not require DEM preprocessing.

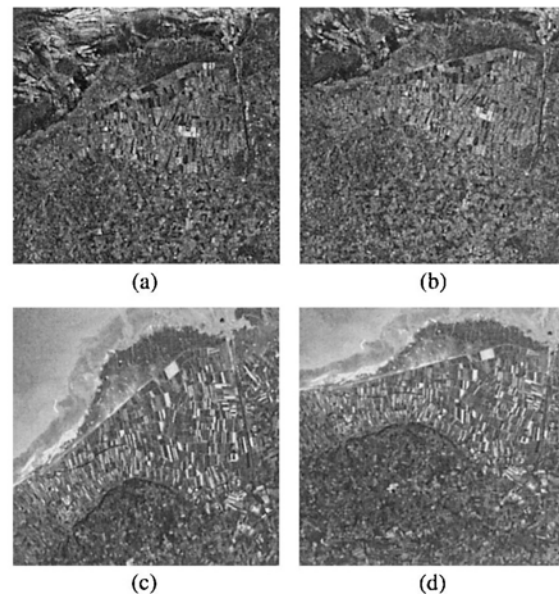


Fig. 5. ERS and SPOT geometrical correction.

The ERS and SPOT image extracts of July (Fig. 5b and d, respectively) are the geometrically corrected versions of the ones in Fig. 5a and c. The geometrical correction processing has not only an effect of rotation but also of scale. A Sun Ultra Sparc 10 takes about 2 min to generate a geometrically corrected extract of 1109 lines \times 1065 columns from a 8200 \times 8000 ERS image or from a 6000 \times 6000 SPOT image.

4.2. Multisensor crosscorrelation

We dispose of four ERS and two SPOT scenes of the same region that necessitate translation error compensation after their geometrical correction. Twenty common points have been extracted from the six images to quantify the reliability of the automatic method of multirate and multisensor registration. In Table 1, are

Table 1
RMSE of the possible registration combinations

Registration type	Combinations	RMSE (pixel)
Multirate ERS	6	0.715
Multirate SPOT	1	0.380
Multisensor	8	0.734

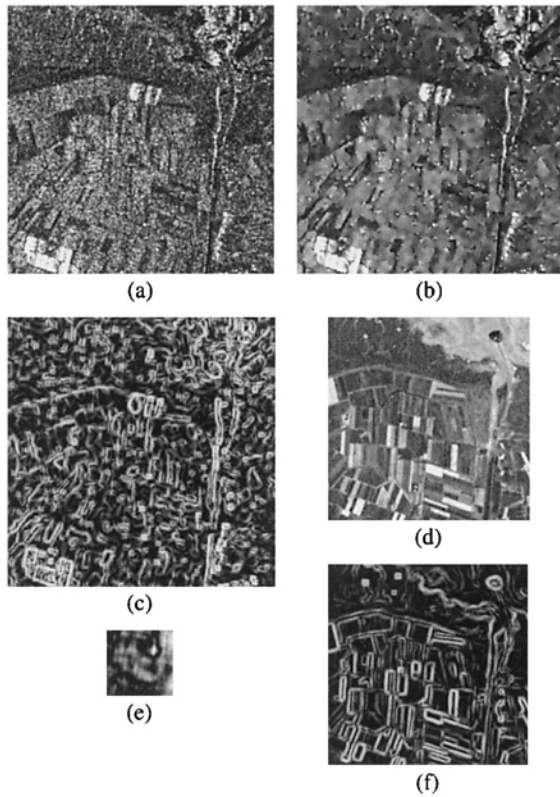


Fig. 6. ERS R and SPOT R crosscorrelation.

presented the results in term of root mean square error (RMSE) of every possible registration combinations of these six images.

The registration quality depends highly of the extract size. The larger they are, the more reliable is the peak location. Nevertheless, we have to make a compromise between this size and the computation time that is about 1 min for the crosscorrelation of a 300×300 extract on a 400×400 one.

Fig. 6e and Fig. 7 present the pixel-by-pixel cross-correlation result of a 300×300 SPOT extract (Fig. 6d) after RMR detection (Fig. 6f) on a 400×400 ERS extract (Fig. 6a) after adaptive Lee filtering (Fig. 6b) and RMR detection (Fig. 6c). As can be seen in Fig. 7, the peak of ρ/ρ_{\max} is easily detectable. Its location is representative of the deviation correction in row and column necessary to register the SPOT image to the ERS one (Fig. 8).

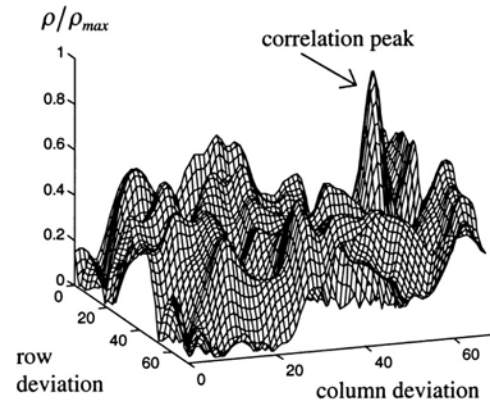


Fig. 7. Extract of Fig. 6e in three dimensions.

4.3. Immediate application

The examples of an immediate application presented in Fig. 9, concern the filtering of speckle, the multiplicative noise inherent in SAR images. From four coregistered ERS images, the linear least mean square error (LMSE) multichannel Lee filter (Lee et al., 1991) and the least-variance multichannel filter (Bruniquel and Lopes, 1997) generate two new images (Fig. 9e and f, respectively) by using the August, September and October (Fig. 9b, c and d, respectively) images to filter the July one (Fig. 9a).

From the four channels and according to the notation of Eqs. (1) and (2), the unspeckled estimated reflectivity is defined by

$$\hat{r}_j = \sum_{i=1}^4 \alpha_i I_i \frac{\mu_j}{\mu_i}$$

The constraints on the weighting coefficients α_j are given respectively for multichannel Lee and Bruniquel filters by Eqs. (5) and (6) from the intensity cross-correlation $\rho_{i,j}$ between i and j channels (Table 2) and the local coefficients of variation C_i

Table 2
Crosscorrelation coefficients

$\rho_{i,j}$	August	September	October
July	0.987	0.980	0.987
August		0.985	0.975
September			0.960

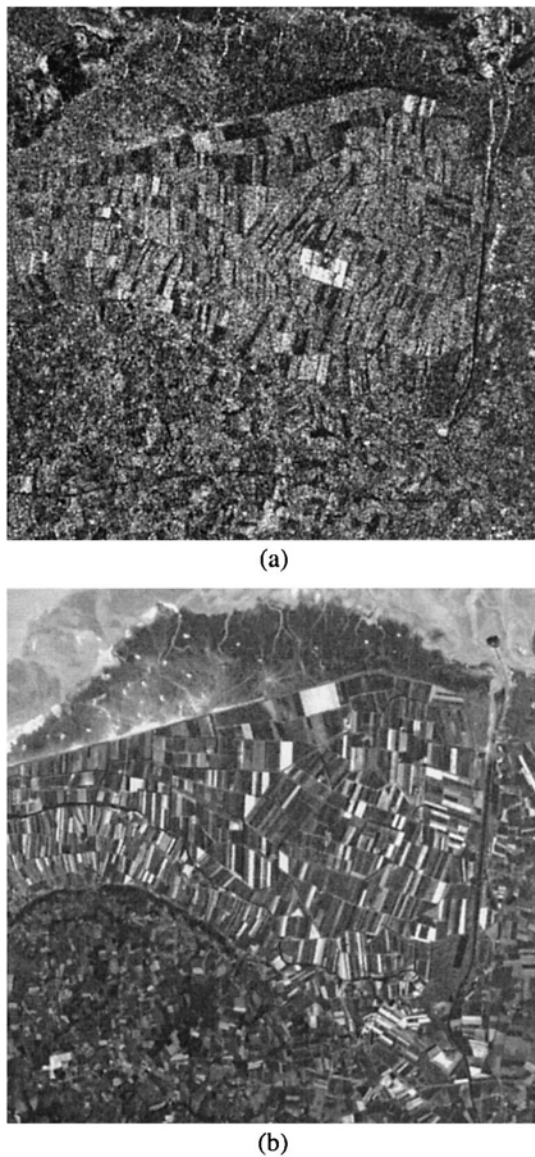


Fig. 8. ERS and SPOT images after translation compensation.

$$\begin{cases} \sum_{i=1}^4 \alpha_i = 1 \\ \sum_{i=1}^4 \alpha_i (\rho_{i,j} - \rho_{i,1}) = 0 \end{cases} \quad j = 2, 3, 4 \quad (5)$$

$$\begin{cases} \sum_{i=1}^4 \alpha_i = 1 \\ \sum_{i=1}^4 \alpha_i (\rho_{i,j} C_i C_j - \rho_{i,1} C_i C_1) = 0 \end{cases} \quad j = 2, 3, 4 \quad (6)$$

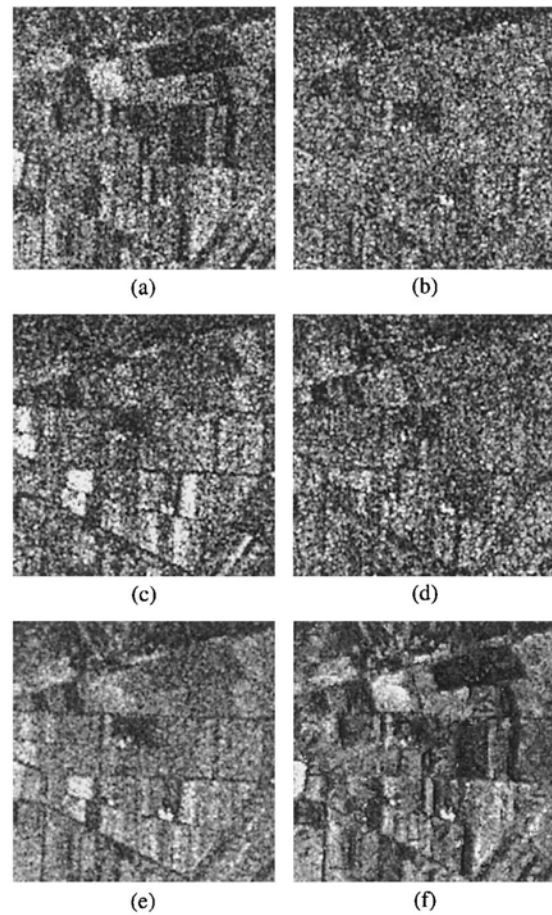


Fig. 9. Lee and Bruniquel multichannel filtering.

As the crosscorrelation coefficient $\rho_{i,j}$, are close to 1, the multichannel Lee filter is equivalent to a simple multichannel average. On the other hand, the Bruniquel one gives better results, because it takes the local variation into account, through the coefficient of variation of the intensity C_i calculated in a 9×9 sub-window.

Moreover, because of the excellent preservation or reconstruction of the field edges and of the roads, the filtering quality of the multichannel Bruniquel filter is a good test of the registration reliability.

5. Conclusion

The data fusion of remote sensing images would seem to be worthwhile considering the evolution of the

natural and agricultural resource management. Unfortunately, the multirate and multisensor images are not directly superposable, so they have to be coregistered.

Here we propose a fully automatic registration method divided into two steps. Geometric correction is followed by residual translation compensation through spatial crosscorrelation processing. This method is efficient for multirate images, but has to be completed for multisensor ones. Edge images, extracted from prefiltered ERS and SPOT images, undergo the same processing as the multirate ones. This reliable method constitutes the unavoidable preliminary step of a multichannel consensual segmentation or classification.

References

- Baraldi, A., Parmiggiani, F., 1995. An alternative form of the Lee filter for speckle suppression in SAR images. *Graph. Models Image Proc.* 57(1), 75–78.
- Bonn, F., Rochon, G., Précis de télédétection: principes et méthodes. In: Presses de l'Université du Québec, vol. 1, Canada, 1992.
- Brown, L.G., 1992. A survey of image registration techniques. *ACM Comput. Surveys* 24(4), 325–376.
- Bruniquel, J., Lopes, A., 1997. Multivariate optimal speckle reduction in SAR imagery. *Int. J. Remote Sens.* 18(3), 603–627.
- Curlander, J.C., McDonough, R.N., Geometric calibration of SAR data. In: *Synthetic Aperture Radar: Systems and Signal Processing*, Wiley, New York, 1991.
- Djardji, J.P., Bijaoui, A., Maniere, R., 1993. Geometrical registration of images: the multiresolution approach. *Photogrammetric Eng. Remote Sens.* 59(5), 645–653.
- Lee, J.S., 1980. Digital image enhancement and noise filtering by use of local statistics. *IEEE Trans. Pattern Anal. Machine Intell.* 2(2), 165–168.
- Lee, J.S., Grunes, M.R., Mango, S.A., 1991. Speckle reduction in multipolarization, multifrequency SAR imagery. *IEEE Trans. Geosci. Remote Sens.* 29(4), 535–544.
- Lopès, A., Nezri, E., Touzi, R., Laur, H., 1993. Structure detection and statistical adaptive speckle filtering in SAR images. *Int. J. Remote Sens.* 14(9), 1735–1758.
- Rignot, E.J.M., Kowk, R., Curlander, J.C., Pang, S.S., 1991. Automated multisensor registration: requirements and techniques. *Photogrammetric Eng. Remote Sens.* 57(8), 1029–1038.
- Thépaut, O., Kpalma, K., Ronsin, J., ERS SAR and SPOT image automatic registration in a multichannel consensual segmentation scheme. *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, vol. 2, pp. 1040–1042, July 1998.
- Touzi, R., Lopès, A., Bousquet, P., 1988. A statistical and geometrical edge detector for SAR image. *IEEE Trans. Geosci. Remote Sens.* 26(6), 764–773.

XVI - Recalage d'images de télédétection

Youcef Bentoutou, Nasreddine Taleb, Kidiyo Kpalma and Joseph Ronsin,

An Automatic Image Registration for Applications in Remote Sensing

IEEE Transactions on Geoscience and Remote Sensing,

September 2005

Vol. 43, No. 9, pp 2127-2137 (11 pages)

Résumé

Cet article traite d'un problème majeur en télédétection concernant le recalage des images multi-temporelles et/ou multicapteurs. En général, de telles images ont différentes caractéristiques de niveau de gris et les techniques simples comme celles basées sur la corrélation ne peuvent pas être appliquées directement. Dans ce travail, nous proposons une nouvelle approche de recalage automatique d'images satellitaires. Cette technique exploite les propriétés invariantes entre les régions d'une image de référence et celle d'une image test. Elle consiste à sélectionner des points de contrôle pertinents basés sur les contours dans l'image de référence. La recherche des points de contrôle homologues dans l'image test est basée sur la détection de similitude locale selon une mesure d'invariants. La superposition finale des images est ensuite réalisée en utilisant l'interpolation par "thin-plate spline". Le processus est entièrement automatique et efficace. L'algorithme proposé pour cette technique a été appliqué avec succès pour recalcr les images SPOT et les images SAR de zones urbaines et agricoles. Les résultats expérimentaux démontrent la robustesse, l'efficacité et la pertinence de l'algorithme.

An Automatic Image Registration for Applications in Remote Sensing

Youcef Bentoutou, Nasreddine Taleb, Kidiyo Kpalma, and Joseph Ronsin

Abstract—This paper deals with a major problem encountered in the area of remote sensing consisting of the registration of multitemporal and/or multisensor images. In general, such images have different gray-level characteristics, and simple techniques such as those based on correlation cannot be applied directly. In this work, a new automatic satellite image registration approach is proposed. This technique exploits the invariant relations between regions of a reference and a sensed image, respectively. It involves an edge-based selection of the most distinctive control points (CPs) in the reference image. The search for the corresponding CPs in the sensed image is based on local similarity detection by means of template matching according to a combined invariants-based similarity measure. The final warping of the images according to the selected CPs is performed by using the thin-plate spline interpolation. The procedure is fully automatic and computationally efficient. The proposed algorithm for this technique has been successfully applied to register multitemporal SPOT and synthetic aperture radar images from urban and agricultural areas. The experimental results demonstrate the robustness, efficiency and accuracy of the algorithm.

Index Terms—Feature extraction, image matching, image registration, remote sensing.

I. INTRODUCTION

IMAGE registration is a classical problem encountered in many image processing applications where it is necessary to perform joint analysis of two or more images of the same scene acquired by different sensors, or images taken by the same sensor but at different times. Examples of these applications include change detection using multiple images acquired at different times [1], cartography and photogrammetry using imagery with overlapping coverage, and fusion of image data from multiple sensor types (e.g., low-resolution multispectral image and high-resolution panchromatic one) [2].

Given two images, I_1 (defined as a reference image) and I_2 (defined as a sensed image) to match the reference image, the goal of image registration is to rectify the sensed image into the coordinate system of the reference image and to make corresponding coordinate points in the two images fit the same geographical location. These unregistered images may have relative translation, rotation, scale between them; they may be degraded

by a linear shift-invariant blur and corrupted by an additive random noise.

In remote sensing applications, one generally uses manual registration, which is not feasible in cases where a large number of images need to be registered. Thus, there is a need for automated techniques that require little or no operator supervision. The automatic image registration addresses the problems associated with manual image registration. However, there still exist a number of scenarios where automatic image registration is not well developed for multitemporal and/or multisensor image registration. The traditional procedure for manually registering a pair of satellite images requires the manual selection of control points in each image. These points are used to determine the parameters of a transformation function, which is subsequently used to register the sensed image to the reference one, by warping one of the images with respect to the other using any interpolation function. Automation of this procedure requires the replacement of the manual control point selection with automatic algorithms for locating corresponding points in both images [3].

The registration process is usually carried out in four steps. The first step consists of selection of features on the images. Next, each feature in one image is compared with potential corresponding features in the other one. A pair of points with similar attributes is accepted as matches and they are called control points (CPs). Finally the parameters of the best transformation which models the deformation between both images are estimated using the CPs obtained in the previous step.

A. Current Approaches

An image registration algorithm suitable for remote sensing use has to meet several criteria.

- 1) It should be accurate.
- 2) It should be reliable. Hence, it should be able to handle all the occurring displacements and gray-level variations between the two images.
- 3) Finally, after fulfilling the above two conditions, the registration should be as fast as possible and ideally fully automatic.

Several image registration techniques have been developed [3], [4]. Unfortunately, none of these techniques has been shown to satisfy all of the above criteria. A review of recent as well as classic image registration methods has been surveyed in [4]. Current image registration techniques are generally divided into two broad categories: area-based and feature-based techniques.

Image registration with the most developed class of area-based techniques promises to be fully automatic [5], [6]. It works correctly when the input images are very similar

Manuscript received February 6, 2004; revised January 10, 2005.

Y. Bentoutou is with the National Center of Space Techniques (CNTS), Arzew, 31200, Algeria (e-mail: bentyou@cnts.dz).

N. Taleb is with the Department of Electronics, University of Sidi-Bel-Abbes, 22000, Algeria (e-mail: ne_taleb@univ-sba.dz).

K. Kpalma and J. Ronsin are with the Department of Electronics and Industrial Informatics, Institut National des Sciences Appliquées de Rennes, 35043 Rennes Cedex, France (e-mail: kidiyo.kpalma@insa-rennes.fr, ronsin@insa-rennes.fr).

Digital Object Identifier 10.1109/TGRS.2005.853187

0196-2892/\$20.00 © 2005 IEEE

but frequently leads to misregistration when dissimilarities are present in the images. In this case, these methods do not satisfy the second constraint. For most of these techniques, some function is defined that quantifies the similarity between the image pairs. The parameter set for which this function is optimal is then taken as the registration transform. The search for this optimum is time consuming since it often requires many function evaluations. Hence, these area-based techniques tend to be somewhat slow. Commonly used area-based similarity functions include correlation, the mean square difference of the image intensity values, and mutual information (MI). The normalized cross-correlation measure is one of the most common similarity functions used in registration. The centers of the matched windows are used as CPs to solve for the transformation parameters between the two images. The correlation measures perform well when the two images to be matched are similar but they become unreliable when the gray-level characteristics of the images are quite different. Mutual information is one of the few intensity-based measures that are well suited to registration of multimodal images [7]–[9]. Unlike measures based on correlation of gray values or differences of gray values, mutual information does not assume a linear relationship between the gray values in the images. The registration algorithm proposed by Thevenaz and Unser in [9] is based on a combination of mutual information together with a multiresolution gradient search. They developed a new optimizer, which takes benefit of the Marquardt–Levenberg strategy, for solving the problem of intermodal image registration, and which works in conjunction with a high-quality multiresolution representation of the image based on cubic splines. Mutual information (MI) has been studied for the registration of remote sensing imagery [10]–[13], and it has been found to be especially robust for multisensor image registration. The transformation which maximizes MI has been determined using exhaustive search [10]–[12], but this has been found to be inefficient and computationally expensive [13]. A multiresolution registration technique proposed recently in [13] combined a powerful search strategy based on a stochastic gradient with two similarity measures, correlation and mutual information, together with a wavelet-based multiresolution pyramid. In this recent work, mutual information has been shown to be better suited for subpixel registration as it produces consistently sharper optimum peaks than correlation [13]. This approach has been applied to registering images with rotational and/or translational differences rather than the natural remotely sensed images with general affine distortion. Affine distortion appears in remote sensing very often; it describes for example the image skew caused by the Earth's rotation [14].

In contrast to the area-based methods, the feature-based ones do not work directly with image intensity values [5], [14]–[22]. Feature-based methods, which extract and match salient features from two images, have been shown to be more suitable for situations when illumination changes are expected or multisensor analysis is performed [5], [20]. There are two critical procedures generally involved in the feature-based techniques: feature detection and feature matching [21]. The first procedure consists of detecting manually or, preferably, automatically a set of distinctive objects such as closed-boundary re-

gions, edges, contours, line intersections, and corners. A variety of image segmentation and edge detection techniques have been used for feature extraction [4]. For further processing, these features can be represented by their point representatives (centers of gravity of closed boundaries [14], [16], [21] and distinctive points [5]), which are called control points in the literature [4]. Feature matching is one of the most important tasks in automated image registration. In this procedure, the correspondence between the features detected in the sensed image and those detected in the reference image is established. Various feature descriptors and similarity measures, such as binary correlation, chain-code correlation, distance of invariant moments, structural matching, Chamfer matching, dynamic programming, and relaxation algorithms, are used for that purpose [4]. A drawback of the existing feature-based techniques lies in the difficulty of recognizing matched features in the images.

Satellite image registration with the most recently developed model incorporating multiple feature extraction algorithms promises to be accurate [21], [22]. However, in both images identical CPs have to be detected manually, which is time-consuming for an operator. Moreover, some extra time has to be added for multiple feature extraction. Consequently, multiple feature extraction and the necessary human supervision make the registration slow and far from fully automatic.

B. Aim of the Paper

The purpose of this paper is to demonstrate that using a gradient-based control point selection mechanism and using the combined invariants-based similarity measure for CP displacement computation, lead to a registration algorithm that satisfies all the constraints mentioned at the beginning of this section and thus can efficiently improve the quality of the registered images. The proposed registration method is performed by applying the elastic thin-plate spline (TPS) interpolating registration on sets of CP pairs [23]. This interpolation model plays an essential role to reach subpixel accuracy. The designed algorithm has the capability to register images both globally and locally. This algorithm, introducing local processing on areas of interest based on control pixels, is fast compared to all traditional algorithms where entire images are searched, and as will be seen, demonstrates the capability to really register satellite images with accuracy. The proposed algorithm, using the TPS interpolation, has been designed to be applied on satellite images as well as on low-altitude aerial pictures of rough terrain.

The rest of the paper is organized as follows. The proposed registration approach is described in detail in Section II. Results of experiments on real data sets are reported in Section III. Finally, the related conclusions are given in Section IV.

II. REGISTRATION APPROACH

In this paper, the registration process is carried out in the following three steps.

- 1) The first step consists of selecting automatically a set of control points $P = \{p_i\}$ from the reference image.
- 2) In the second step, the displacements of the selected control points $p_i \in P$ are computed by means of template matching using the combined invariants-based similarity



Fig. 1. Summary of the proposed automatic image registration method.

measure. The corresponding coordinates of the selected CPs in the sensed image constitute the set of CPs $Q = \{q_i\}$ in this image.

- 3) Last, the parameters of the best transformation which models the deformation between the images are estimated using the two sets of corresponding control points (P and Q). Then the final correction is performed according to this transformation by warping one of the images with respect to the other one using the TPS interpolation function and the estimated parameters [23].

This proposed registration approach is summarized in Fig. 1.

A. Control Points Selection

In order to obtain the correspondence between two images, area-based techniques would explicitly compute the correspondence for every pixel. However, this is computationally very expensive and will not lead to practically acceptable algorithms. In order to reduce the computation time to an acceptable level, a frequently used approach is to consider only a limited set of so-called interest points, where the image information is supposed to be the most important [24]. It has to be noted that in the particular case of satellite images, misalignment between reference and sensed images is more visible at the image edges than at the homogeneous regions. Therefore one easy way to speed up the algorithm and to reduce the computation time is to process some few pixels belonging to regions that contain significant gray-level variations (referred to as the control points), instead of processing all image edges. Obviously, the less complex the image, the faster the algorithm [19].

1) *Edge Detection*: The first step in the selection of the CPs consists of finding image edges. The location of edges can easily be computed by detecting local maxima of the gray-level gradient magnitude. Since edges are scale dependent image features, they can only be detected properly using derivatives calculated at the proper scale. The commonly used edge detection filter for this purpose is Gaussian derivative that allows finer tuning to detect the required edge scale [25]. It should be noted that it is necessary to indicate which edges are important for extracting adequate CPs. This is done by thresholding the gradient magnitude image at a value T_G , resulting in a binary image consisting of connected regions that are of interest. The value of the threshold T_G is assigned as the average gradient magnitude [26].

2) *Control Points Detection*: To select CPs in the reference image, the second step consists of detecting interest points that have stronger strength than most of other pixels in this image. The strength is basically measured by the change in pixel values along two-dimensional directions. Schmid *et al.* [26] evaluated and compared different interest point detectors and concluded that the best results are provided by the Harris detector [27]. In this paper, an improved version of the Harris corner detector [26] is used to detect points of interest in the reference image. This detector uses only first-order derivatives and is well known

as one of the most stable and robust corner detectors in image processing. The basic idea is to use the autocorrelation function in order to determine locations where the signal changes in two directions. A matrix related to the autocorrelation function that takes into account the first derivatives of the signal on a window is computed. The eigenvectors of this matrix are the principal curvatures of the autocorrelation function. Two significant values indicate the presence of an interest point. The Harris detector is defined as the positive local extreme of the following operator:

$$S = \det(C) - \alpha \text{trace}^2(C) \quad (1)$$

where scalar α is equal to 0.04 as suggested in [27] and the matrix C , which is related to the autocorrelation function, is given by

$$C = \begin{bmatrix} \left(\frac{\partial I}{\partial x}\right)^2 & \left(\frac{\partial I}{\partial x}\right) \cdot \left(\frac{\partial I}{\partial y}\right) \\ \left(\frac{\partial I}{\partial x}\right) \cdot \left(\frac{\partial I}{\partial y}\right) & \left(\frac{\partial I}{\partial y}\right)^2 \end{bmatrix} \quad (2)$$

where I is the gray-level intensity image. To implement this detector, the standard default parameter values recommended by the authors, are used in this work. The Harris detector [27] computes the derivatives of the matrix C by convolution with the mask $[-2 \ -1 \ 0 \ 1 \ 2]$. A Gaussian with standard deviation ($\sigma = 2$) is used to weight the derivatives. In an improved version of Harris [26], the derivatives of the matrix C are computed more precisely by replacing the $[-2 \ -1 \ 0 \ 1 \ 2]$ mask with derivatives of a Gaussian ($\sigma = 1$). A corner is detected when the two eigenvalues of C are large since this indicates grayscale variations in both the x and y directions. To avoid the extraction of the eigenvalues of the matrix C , the strength of interest points is measured by the operator S of (1). Nonmaximum suppression using a 3×3 mask is then applied to strength and a threshold T is used to select interest points (referred to as the control points). The threshold is set to 1% of the maximum observed interest point strength [26]. To spread out the detected CPs (corners), a user-defined distance is incorporated. Finally, the algorithm will automatically produce the extracted CPs that satisfy the above criteria and maximize the strength function. The number of these CPs is user-defined and incorporated prior to running the detector.

In this implementation, the CPs are detected in the reference image by applying the improved version of the Harris corner detector at regions where the gradient magnitude is high. This has the following two advantages.

- 1) The time complexity is reduced by excluding pixels with low values of gradient magnitude from computation.
- 2) Since the neighborhoods of those interest points are known to be structured, the reliability of the matching step will be high.

3) *Control Points Selection Algorithm*: The following is a summary of the different steps to be followed in the selection of CPs.

- Step 1) Compute the gradient of the reference image. It is based on an edge detection filtering with the application of Gaussian derivative ($\sigma = 1$) [25].

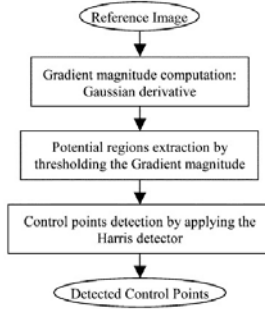


Fig. 2. Gradient-based CP selection mechanism.

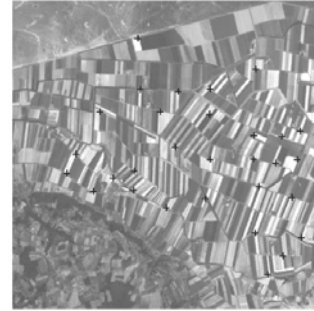
- Step 2) Extract potential regions by means of thresholding the gradient magnitude image at level T_G (T_G = average gradient magnitude) and setting their pixels p to "1".
- Step 3) For each pixel p in the reference image compute the strength S according to (1).
- Step 4) Apply nonmaximum suppression to the strength by using a 3×3 mask.
- Step 5) Extract strong corners by means of thresholding the strength at level T .
- Step 6) Retain the first n corners of large strength as control points if their distance is higher than the minimum allowed distance d_{\min} .

Fig. 2 shows a block diagram of this selection algorithm. An example of CPs selection based on this approach is shown in Fig. 3.

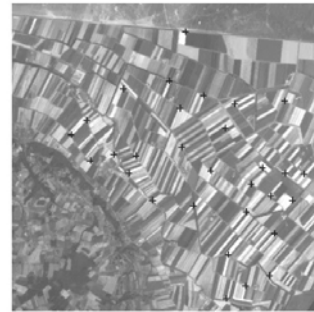
B. Control Points Displacements Computation

After the CPs are detected from the reference image in the previous section, a correspondence mechanism between image regions around these points must be established to match these two images. This correspondence mechanism is called template matching. Template matching techniques are based on the assumption that the displacement d of every CP in an image $I_R(x, y)$ can be approximated by taking a small window W_d around the point and finding the corresponding window in the unregistered image $I_S(x, y)$, according to a certain similarity measure.

The most important aspect of template matching is the similarity measure that is used to determine the degree of resemblance of windows in two images. Several measures have been devised and applied to register images, including correlation, normalized cross-correlation, statistical correlation, matched filters, phase-correlation, sum of absolute differences, root mean square, and masked correlation [3], [4]. However, most of these similarity measures do not satisfy the three constraints that have been emphasized in the introduction of this paper. They are sensitive to mean gray-level offsets and local geometric distortion. In addition to these fundamental problems, there are other factors that may complicate the process of finding the optimal correspondence between images. These are due to imperfections of the imaging system, such as limited spatial resolution, gray-level quantization, blur, and noise.



(a)



(b)

Fig. 3. Control points detection on two SPOT panchromatic images of the region of Mont Saint-Michel in Normandie (France), acquired at different dates. (a) The set of selected CPs (black crosses), with the reference image superimposed. These interest points are detected by applying the improved version of the Harris detector at potential regions where the gradient magnitude is high. The number n of CPs is 30 and the minimum allowed distance d_{\min} between the points is 30 pixels. (b) The corresponding set of the selected CPs (black crosses), with the sensed image superimposed. These corresponding points in the sensed image are detected by means of template matching according to the invariant approach.

1) *Combined Invariants-Based Similarity Measure*: Since the images may have translational, rotational, and scaling differences, the similarity measures should be invariant with respect to translation, rotation, and scaling. Measures that are very robust against these phenomena are those of the moment-based combined invariants (i.e., invariants with respect to degradations composed from symmetric blur and other geometric or gray-level transforms) [28]. Detailed definitions of these combined invariants can be found in literature (see [28]).

The moment of a window W_d inside an image $I(i, j)$ is given by

$$m_{pq} = \sum_{(i,j) \in W_d} i^p j^q I(i, j) \quad (3)$$

where $(p + q)$ is the order of moment.

The central moment of W_d is defined as

$$\mu_{pq} = \sum_{(i,j) \in W_d} (i - x_t)^p (j - y_t)^q I(i, j) \quad (4)$$

where $x_t = (m_{10}/m_{00})$ and $y_t = (m_{01}/m_{00})$ are the coordinates of the center of gravity of W_d . Central moments are translational invariants. Under a scale change $i' = \alpha i$ and $j' = \alpha j$, the normalized central moments are defined as $\nu_{pq} = (\mu_{pq}/\mu_{00}^{(p+q+2)/2})$.

Central moments and scaling-invariant moment representation can be employed to produce a set of invariant moments that are further invariant to symmetric blur and rotation. A set of seven moment invariant functions based on the second- and third-order moments have been first derived by Hu [29]. These functions are invariant to scaling, translation, and rotation. Wong [30] has later proposed a method to generate an infinite sequence of rotation moment invariants and has shown that Hu's invariants are just particular representatives of them. Recently, Flusser and Suk [28] have shown that some odd-order Wong's invariants (or certain simple functions of them) are also blur invariants. Based on this, the set of combined invariants (features that are based on image moments and that are invariant to symmetric blur, scaling, translation, and rotation) used in this work are those introduced in [28] and are given by the following equations:

$$\Phi_1 = (\nu_{30} - 3\nu_{12})^2 + (3\nu_{21} - \nu_{03})^2 \quad (5)$$

$$\Phi_2 = (\nu_{30} + \nu_{12})^2 + (\nu_{21} + \nu_{03})^2 \quad (6)$$

$$\Phi_3 = (\nu_{30} - 3\nu_{12})(\nu_{30} + \nu_{12}) \left((\nu_{30} + \nu_{12})^2 - 3(\nu_{21} + \nu_{03})^2 \right) + (3\nu_{21} - \nu_{03})(\nu_{21} + \nu_{03}) \left(3(\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2 \right) \quad (7)$$

$$\Phi_4 = (3\nu_{21} - \nu_{03})(\nu_{30} + \nu_{12}) \left((\nu_{30} + \nu_{12})^2 - 3(\nu_{21} + \nu_{03})^2 \right) - (\nu_{30} - 3\nu_{12})(\nu_{21} + \nu_{03}) \left(3(\nu_{30} + \nu_{12})^2 - (\nu_{21} + \nu_{03})^2 \right) \quad (8)$$

$$\Phi_5 = [\nu_{50} - 10\nu_{32} + 5\nu_{14} - 10(\nu_{20}\nu_{30} - \nu_{30}\nu_{02} - 3\nu_{12}\nu_{20} + 3\nu_{12}\nu_{02} - 6\nu_{11}\nu_{21} + 2\nu_{11}\nu_{03})]^2 + [\nu_{05} - 10\nu_{23} + 5\nu_{41} - 10(\nu_{02}\nu_{03} - \nu_{03}\nu_{20} - 3\nu_{21}\nu_{02} + 3\nu_{21}\nu_{20} - 6\nu_{11}\nu_{12} + 2\nu_{11}\nu_{30})]^2 \quad (9)$$

It was shown by Flusser and Suk [28] that these combined invariants are sufficiently robust with respect to noise. The local displacement of every CP in the sensed image can be estimated over a circular neighborhood of each one of these points detected earlier. After that, the correspondence is established by the minimum distance rule with thresholding in the Euclidean space of the invariants. The distance between two different circular templates W_1 and W_2 is defined as

$$d_r(w_1, w_2) = \|\vec{C}(r)^{(w_1)} - \vec{C}(r)^{(w_2)}\| \quad (10)$$

where $\|\cdot\|$ is the Euclidean norm in $l_2(r)$ space and the vector $\vec{C}(r)$ of combined invariants is defined as

$$\vec{C}(r) = (\Phi_1, \Phi_2, \dots, \Phi_r) \quad (11)$$

where r is the number of the combined invariants corresponding to the size of $\vec{C}(r)$ (here, $r = 5$). The smaller the distance, the more similar the two regions. In the case of optimal alignment, the distance between the two windows is nil.

2) *Template Matching Algorithm*: The displacements of CPs are determined by carrying out the following steps.

- Step 1) For every control point p_i , choose a moving circular template W_1 of radius ρ centered at this point in the reference image.
- Step 2) Assuming that misregistration parameters would not exceed an " L " pixel shift for image translation, calculate the combined invariants vectors $\vec{C}(r)^{(W_1)}$ and $\vec{C}(r)^{(W_2)}$ of W_1 and the corresponding template W_2 centered at q_j in the sensed image, respectively.
- Step 3) Calculate the distance d_r between the combined invariants vectors of these two templates in the invariant space, taking into consideration the amount of the maximum translational shift. Save the translational shift giving the minimum distance. To simplify the search process, the preliminary false matches between template pairs are excluded if their distances in the Euclidean space of the invariants are above a certain threshold T_d . At the end, the values of the translational shifts will correspond to the displacements of the selected control points up to integer precision.
- Step 4) The control point $q_i \in Q$ in the sensed image that might correspond to the candidate point $p_i \in P$ in the reference image is determined by the shift position giving the minimum distance.

3) *Subpixel Precision*: Many applications like change detection, passive navigation, feature location measurements in remote sensing, and image sequence analysis require registration results with an error less than one pixel (subpixel accuracy) [31]. Therefore, it is necessary to provide means to obtain subpixel precision in the displacements computation because the described matching approach only determines the displacements up to integer precision. Since the image pixels constitute the only available information about the represented scene, this inevitably requires some form of interpolation. In order to obtain subpixel accuracy in the displacements computation, several approaches have been proposed, such as analytical match-interpolation methods and image-interpolation methods [19]. In the image interpolation approach, subpixel accuracy can be achieved by translating the mask image by subpixel increments in the local region around the initial best match position [32]. This method is computationally very expensive compared to the analytical match-interpolation methods. An analytical match-interpolation method [19], [33] is used in this paper. In this case, subpixel accuracy is achieved by interpolating in the match measure surface to estimate the position of the peak. Separate horizontal and vertical quadratic interpolation is used to refine the estimates of the two components of each control point p_i . Given the discrete match measure surface $M(u, v)$, corresponding to the previously defined distance d_r between the template in the sensed image at coordinates (u, v) and the one in the reference

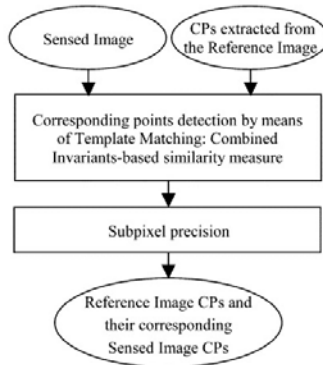


Fig. 4. Block diagram of the CPs displacements computation.

image at the corresponding coordinates (u', v') , and given a best match position (x_0, y_0) , then the estimated continuous coordinates (x, y) of the match peak around (x_0, y_0) are calculated as

$$x = x_0 - \frac{1}{2} + \frac{M(x_0, y_0) - M(x_0 - 1, y_0)}{2M(x_0, y_0) - M(x_0 - 1, y_0) - M(x_0 + 1, y_0)}$$

$$y = y_0 - \frac{1}{2} + \frac{M(x_0, y_0) - M(x_0, y_0 - 1)}{2M(x_0, y_0) - M(x_0, y_0 - 1) - M(x_0, y_0 + 1)} \quad (12)$$

The exact coordinates (x_i, y_i) , in the sensed image of each control point p_i , $i = 1, 2, \dots, n$, constitute the second set of control points $Q = \{q_i\}$ with $q_i = (x_i, y_i)$, $i = 1, 2, \dots, n$. This template matching method is described in Fig. 4.

C. Image Warping

Given the two sets of corresponding control points, $P = \{p_i\}$ and $Q = \{q_i\}$, the transformation parameters are estimated and the correction is performed by warping the sensed image. Satellite image warping can be performed by applying an affine transform [5], [21]. The affine mapping function is appropriate for data with flat topography. For some data with nonlinear, local geometric distortions (such as hilly areas with terrain changes), more complex transformation functions may be needed to produce better interpolation results. These mapping functions include thin-plate spline interpolation. Thin-plate splines are perhaps the most widely used transformation functions in the registration of images with nonlinear geometric differences. They were first used by Goshtasby [15] in the registration of remote sensing images. This interpolation model plays an essential role to reach subpixel accuracy. Other warping functions such as B-splines have been used to achieve subpixel accuracy [9], [13]. In this paper, sensed image warping has been performed by applying the TPS interpolation function [23] on the two sets of corresponding CPs. The proposed algorithm, using the TPS interpolation, has been designed to be applied on satellite images as well as on low-altitude aerial pictures of rough terrain.

The thin-plate splines are interpolating functions, representing, exactly, the distortion at each CP and defining a

minimum-curvature surface between CPs. The TPS function is a flexible transformation that allows rotation, translation, scaling, and skewing. It also allows lines to bend according to the thin-plate spline model. Therefore, a large number of deformations can be characterized by the thin-plate spline model. The thin-plate spline interpolation function can be written as

$$\mathbf{h}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{t} + \sum_{i=1}^n \mathbf{W}_i K(\|\mathbf{x} - \mathbf{x}_i\|) \quad (13)$$

where \mathbf{A} and \mathbf{t} are the affine transformation parameter matrices, \mathbf{W}_i are the weights of the nonlinear radial interpolation function K , and \mathbf{x}_i are the CPs. The function $K(\lambda)$ is the solution of the biharmonic equation ($\Delta^2 K = 0$) that satisfies the condition of bending energy minimization, namely $K(\lambda) = \lambda^2 \log(\lambda^2)$.

To solve for the x, y components, the transformation \mathbf{h} is defined from two sets of n corresponding CP pairs as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} h_x(x, y) \\ h_y(x, y) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} + \left(\sum_{i=1}^n W_{xi} K(\|(x, y) - (x_i, y_i)\|) \right) \quad (14)$$

where (x', y') and (x, y) are two corresponding CPs in the reference image and the sensed image, respectively.

The warping of the sensed image with respect to the reference image is accomplished by carrying out the following steps.

- Step 1) Given the sets P and Q of corresponding coordinates, the estimation of the affine transformation parameters \mathbf{A} and \mathbf{t} is performed by using only the three initial control point pairs having the minimum distance in the Euclidean space of the invariants. These parameters are obtained by solving the following equation in a least square sense:

$$\mathbf{Y} = \mathbf{M}\mathbf{z} \quad (15)$$

where

$$\mathbf{Y} = \begin{bmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ x'_3 \\ y'_3 \end{bmatrix} \quad \mathbf{M} = \begin{bmatrix} x_1 & y_1 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_1 & y_1 & 0 & 1 \\ x_2 & y_2 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_2 & y_2 & 0 & 1 \\ x_3 & y_3 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_3 & y_3 & 0 & 1 \end{bmatrix}$$

$$\mathbf{z} = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ t_x \\ t_y \end{bmatrix}$$

a_{11} , a_{12} , a_{21} , a_{22} , t_x , and t_y are the affine transformation parameters and are found by taking the inverse of the matrix M .

$$z = M^{-1}Y \quad (16)$$

- where M^{-1} is the inverse of M . Based on this transformation, it is possible to map the other control points in the sensed image into the reference image.
- Step 2) The invalid control point pairs are removed during this step. The distance between the mapped control point in the sensed image and the actual point in the reference image is computed. This distance is actually the RMSE at this pair of corresponding points. A threshold can be easily set to avoid false matches. The invalid control point pairs can be removed by examining the following condition:

$$\sqrt{(\hat{x}_i - x'_i)^2 + (\hat{y}_i - y'_i)^2} < E_r \quad (17)$$

where (x'_i, y'_i) and (\hat{x}_i, \hat{y}_i) are the coordinates of the control point pairs in the reference image and the affine transformed sensed image, respectively, and $\hat{x}_i = a_{11}x_i + a_{12}y_i + t_x$ and $\hat{y}_i = a_{21}x_i + a_{22}y_i + t_y$. E_r is a preset threshold that is actually the maximum RMSE allowed at the control points for the final image registration. If this condition is satisfied, the pair of points is kept for further processing (i.e., sensed image resampling). If not, this pair of points is removed from the set of corresponding control points. The output from this step is a set of m corresponding control point pairs.

- Step 3) The two sets of m corresponding control point pairs are then used to calculate the weights W_{x_i} and W_{y_i} of the TPS function using the equations

$$\begin{bmatrix} x'_1 - a_{11}x_1 - a_{12}y_1 - t_x \\ x'_2 - a_{11}x_2 - a_{12}y_2 - t_x \\ \vdots \\ x'_m - a_{11}x_m - a_{12}y_m - t_x \end{bmatrix} = \begin{bmatrix} 0 & K(\lambda_{12}) & \dots & K(\lambda_{1m}) \\ K(\lambda_{21}) & 0 & \dots & K(\lambda_{2m}) \\ \vdots & \vdots & \ddots & \vdots \\ K(\lambda_{m1}) & K(\lambda_{m2}) & \dots & 0 \end{bmatrix} \begin{bmatrix} W_{x1} \\ W_{x2} \\ \vdots \\ W_{xm} \end{bmatrix} \quad (18)$$

$$\begin{bmatrix} y'_1 - a_{21}x_1 - a_{22}y_1 - t_y \\ y'_2 - a_{21}x_2 - a_{22}y_2 - t_y \\ \vdots \\ y'_m - a_{21}x_m - a_{22}y_m - t_y \end{bmatrix} = \begin{bmatrix} 0 & K(\lambda_{12}) & \dots & K(\lambda_{1m}) \\ K(\lambda_{21}) & 0 & \dots & K(\lambda_{2m}) \\ \vdots & \vdots & \ddots & \vdots \\ K(\lambda_{m1}) & K(\lambda_{m2}) & \dots & 0 \end{bmatrix} \begin{bmatrix} W_{y1} \\ W_{y2} \\ \vdots \\ W_{ym} \end{bmatrix} \quad (19)$$

where $\lambda_{i,j} = \|(x_i, y_i) - (x_j, y_j)\|$, $1 \leq i \leq m$, $1 \leq j \leq m$, and (x_i, y_i) and (x_j, y_j) are the coordinates of the m control points in the sensed image.

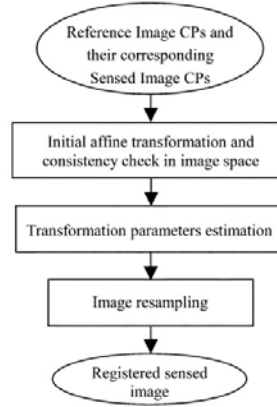


Fig. 5. Schematic diagram for the sensed image warping method.

- Step 4) The complete set of parameters, defining the interpolating registration transformation is then used to transform the sensed image values. It should be noted that in order to be able to carry out the warping of the sensed image with respect to the reference image, it is required to have a complete description of the TPS interpolation function (14). That is, this function must be known for every point in the image. So far, only the method for computing the TPS function of a selected set of control points has been described in this paper, under the assumption that the remainder of the field could be obtained by linear interpolation. Detailed algebraic treatment of this TPS-based warping can be found in literature (see [23]).

This warping method is summarized in Fig. 5.

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

To measure the performance of the proposed registration approach, the experimental results were divided into two parts. First, in order to evaluate the accuracy of the registration approach, the proposed method was tested by using a pair of test images: the sensed image being the rotated version of the reference one (see Fig. 6). Second, the proposed method was applied for the registration of multitemporal satellite images from Systeme Pour l'Observation de la Terre (SPOT) and European Remote Sensing 1 synthetic aperture radar (SAR) sensors. In this latter experiment, two image pairs of the region of Mont Saint-Michel in Normandie (France) were registered separately, one pair of SPOT images and the other one of SAR images acquired at different dates. Subscenes of around 512×512 pixels from the original images were used. The parameters of the Harris detector are set to the standard default values recommended by the authors [26]. The standard deviation of the Gaussian is set to 1 for all examples. During these experiments, the minimum allowed distance between two CPs and the desired number of CPs are set to $d_{\min} = 30$ and $n = 10$. The maximum translational shift used in template matching is set to 100 and the radius ρ of the circular templates is 20. There is no simple

criterion for the template radius ρ . In order to reduce the computation time to a minimum, this template should be as small as possible. However, in principle, the size of the template determines the amount of statistical information that is provided. Small templates will yield unreliable match values. Experiments in this work have indicated that a template of radius size between 10 and 20 pixels yields a good compromise between computational speed and statistical reliability. To identify the most robust matches, the distance threshold in the Euclidean space of the invariants T_d is set to 0.1 and the maximum RMSE (E_r) allowed at the CPs is set to 0.3.

The registration results of this method were compared with those of manual methods. In the existing methods, the root mean square error (RMSE) between the matched CPs provides a measure of registration (20). The RMSE at m matched CPs (in pixels) is defined as in (20), shown at the bottom of the page.

A. Comparison of the Proposed Affine Registration Results With Those of the Manual Registration

For a good evaluation of the performance of the proposed algorithm, distinct corner points were manually selected as CPs on two SPOT images, a reference image and its rotated version as the sensed image (see Fig. 6). The transformation parameters and RMSEs obtained by manual CP selection and registration were compared with those of the proposed automatic CP selection and affine registration with no TPS (see Table I). Hence, the RMSEs in this case have been obtained for both the manual and automatic methods according to (20) using just the affine parameters without the terms related to the TPS function. Ten pairs of CPs are manually selected to register the two images. The number of manually selected CPs and the number of CPs generated by the proposed algorithm are kept the same in order to compare their RMSEs. Thus, to apply the automatic proposed method, the ten selected CPs in the reference image were kept the same as those of the manual method in order to automatically generate the corresponding points on the sensed image. From Table I, it is clear that the differences between the transformation parameters of the manual registration method and that of the proposed one are considerable. The RMSE difference between them is greater than two pixels. The transformation parameters obtained by the proposed registration approach are very close to the optimum values (i.e., the true/known transformation parameters). Clearly the image registration, using the proposed algorithm, results in a lower RMSE, and its accuracy is therefore better than that of the manual one.

B. Registration of Multitemporal SPOT and SAR Images

The results of the application of the proposed registration technique to a pair of SPOT and SAR images acquired at dif-

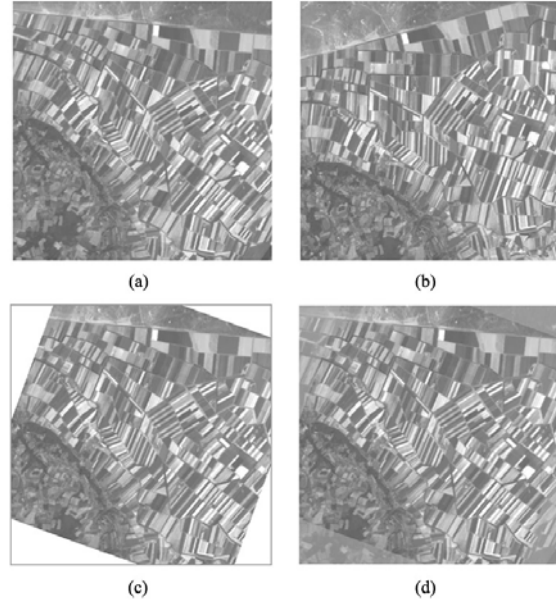


Fig. 6. Registration of a pair of test images. (a) Initial SPOT image. (b) Rotated image by a 10° clockwise rotation of the initial image. The center of rotation is the center of the image. (c) Image (b) after geometric transformation. (d) The registration result using the proposed approach.

TABLE I
COMPARISON OF THE PROPOSED AUTOMATIC REGISTRATION RESULTS WITH THE MANUAL REGISTRATION RESULTS AND WITH GROUND TRUTH

Methods	a_{11}	a_{12}	a_{21}	a_{22}	t_x	t_y	RMSE
Optimal	0.9396	0.3420	-0.3420	0.9396	0.0000	0.0000	0.0000
Manual	0.9450	0.3501	-0.3440	0.9470	0.9503	1.5602	2.8162
Proposed Method	0.9395	0.3422	-0.3419	0.9397	0.0504	3.2e-05	0.0385

ferent dates are shown in Figs. 7 and 8. In these figures, four images are chosen for each pair: (a) the reference image, (b) the sensed image, (c) the registered sensed image, and (d) the registration result of the two images after registration. Based on the algorithm parameters set, there are ten CPs in the reference image that were automatically detected by the improved Harris corner detector. The corresponding CPs in the sensed image are computed by means of template matching using the combined invariants-based similarity measure. The coordinates of all CPs determined after application of the image matching step of the algorithm are summarized in Table II.

Given a number of matched corresponding control points from two images, the affine transformation parameters can be estimated. The registration of images with rotational, translational, scaling, and image-skew differences can be approximated by applying the mapping functions of (14). The

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^m \left(a_{11}x_i + a_{12}y_i + t_x + \sum_{j=1}^m W_{xj}K(\lambda_{ij}) - x'_i \right)^2 + \left(a_{21}x_i + a_{22}y_i + t_y + \sum_{j=1}^m W_{yj}K(\lambda_{ij}) - y'_i \right)^2}{m}} \quad (20)$$

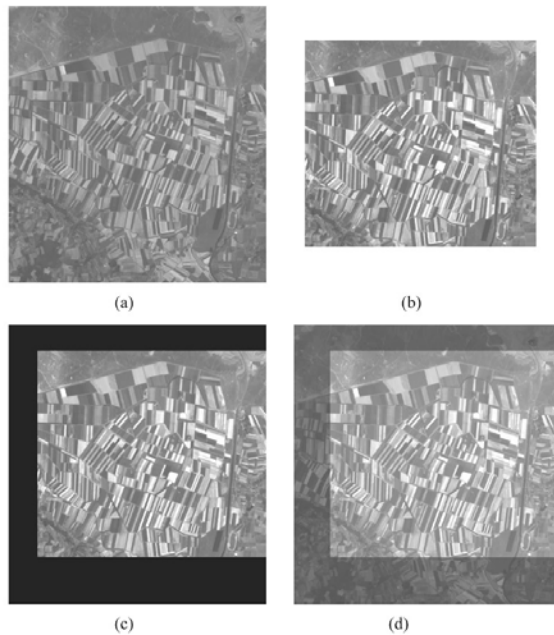


Fig. 7. Registration of SPOT panchromatic images of the region of Mont Saint-Michel in Normandie (France), acquired at different dates. (a) The reference image. (b) The sensed image. (c) The registered sensed image obtained after application of the proposed approach. (d) The registration result.

thin plate-spline transformation function is appropriate for satellite images with nonlinear, local geometric distortions and is needed to produce better interpolation results and to achieve subpixel accuracy. For this type of transformation, the affine transformation parameters and the weights of the nonlinear radial interpolation function are computed in order to resample and transform the sensed image according to the reference image. Since the determination of the affine transformation parameters requires the knowledge of only three CPs, three pairs of CPs with minimum distance were detected and used to estimate these parameters. Based on this affine transformation, the invalid CP pairs (i.e., false matches) are removed if the RMS error (E_r) exceeds a specified threshold as shown in (17). In this case, seven additional matched points were detected in the pair of SPOT and SAR images. An RMSE of less than 0.3 pixels has been obtained for the ten CPs. Knowing the TPS mapping function, the sensed image was transformed and resampled using linear interpolation. The results of the proposed approach are shown in Figs. 7(c) and 8(c).

The registration accuracy of the proposed method was estimated using the RMSE of (20) at every CP, as shown in Table III. RMSE1 denotes the errors after registering two multitemporal SPOT images, while RMSE2 denotes the errors after registering two multitemporal SAR images. The results show that in the proposed automated algorithm, most errors (RMSEs) are less than 0.3 pixels. In this approach, the estimation of the affine transformation parameters is performed by using only the three CPs having the minimum distance among the ten (control points 1, 2, and 3; see Table III). This has the following two advantages.

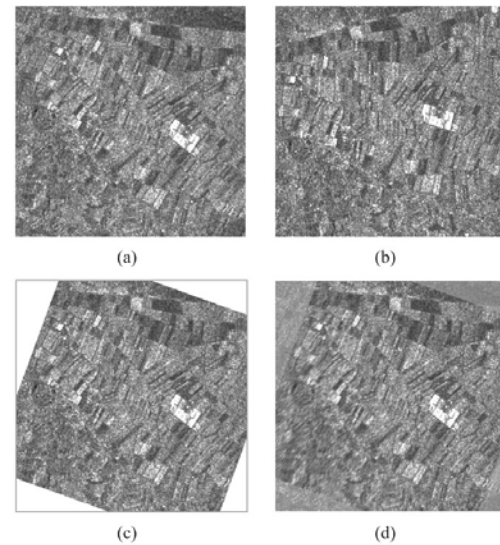


Fig. 8. Registration of SAR images of the region of Mont Saint-Michel in Normandie (France), acquired at different dates. (a) The reference image. (b) The sensed image. (c) The registered sensed image obtained after application of the proposed approach. (d) The registration result.

TABLE II
COORDINATES OF CORRESPONDING CONTROL POINTS
IN THE TWO PAIRS OF IMAGES

Control Point	SPOT Images		SAR Images	
	Reference (x, y)	Sensed (x', y')	Reference (x, y)	Sensed (x', y')
1	(450, 453)	(533.99, 513.96)	(285, 48)	(369.98, 75.94)
2	(140, 96)	(223.97, 156.98)	(385, 84)	(452.16, 143.92)
3	(38, 189)	(122.01, 250.02)	(416, 146)	(459.95, 212.90)
4	(37, 324)	(120.88, 384.98)	(398, 274)	(399.10, 327.12)
5	(113, 439)	(197.08, 499.88)	(364, 417)	(316.88, 451.07)
6	(206, 467)	(290.07, 527.77)	(236, 434)	(188.99, 422.96)
7	(378, 437)	(461.97, 498.11)	(73, 302)	(81.12, 243.03)
8	(506, 358)	(589.88, 419.02)	(38, 167)	(94.98, 103.97)
9	(475, 249)	(559.01, 309.97)	(116, 134)	(180.12, 99.07)
10	(498, 149)	(582.10, 210.01)	(277, 147)	(328.09, 166.10)

TABLE III
ROOT MEAN SQUARE ERRORS AT THE CONTROL POINTS (IN PIXELS)

CP	1	2	3	4	5	6	7	8	9	10	Total CPs RMSE
RMSE1	0.0252	0.0061	0.0067	0.1698	0.1717	0.2679	0.1173	0.0977	0.0806	0.2119	0.1433
RMSE2	0.0087	0.0122	0.0118	0.2272	0.2253	0.2208	0.2559	0.1888	0.0705	0.0288	0.1607

- 1) The computation time is reduced with respect to processing all the detected control points.
- 2) The use of only three control points with minimum distance enhance more the accuracy of the affine transformation parameters estimation and thus the effectiveness of the registration than that using the entire ten points some of them having no minimum distance. The total control points RMSE in this case is around 0.1 pixels, which is a very high accuracy.

In general, image registration, using the proposed algorithm, results in a lower RMSE.

The proposed algorithm has several user-defined parameters, namely the number n of CPs, the minimum allowed distance

d_{\min} between the points, the radius ρ of the circular template, and the maximum translational shift L used for template matching. Increasing the parameters ρ , L , and n slows down the registration but may enhance the accuracy and robustness. When the geometric deformations between the images are large and fast registration is needed, applying a coarse rigid preregistration before extracting the corresponding points is recommended. This will reduce the required maximum translational shift used in template matching and, consequently, increase the speed of registration. The choice of the parameters is influenced by the type of the scene. The proposed algorithm was successfully tested on one scene. In the case of highly textured scenes where features may not be reliably observed, a large number of CPs can be selected corresponding to reference image corners or highly textured patches. Since the estimation of the transformation parameters is sensitive to location errors when the features are close to each other, it is convenient to spread out the detected CPs by setting the minimum allowed distance parameter d_{\min} between two CPs high. The template matching stage of the algorithm [i.e., Steps 1)–4)] of Section II-B2 may generate a false match, for example, due to repetitive texture patterns. In this case, the use of large templates for feature matching allows for a reliable correspondence between the images. This algorithm is quite robust and reliable as long as corresponding distinctive points are available in the images. The registration scheme would fail if the input images did not contain distinctive features, e.g., the case of mountainous scenes without any distinctive features. However, in such situation, almost every feature-based matching procedure will fail.

IV. CONCLUSION

In this paper, a new approach to the registration of remotely sensed data is proposed and tested on one scene. In this approach, a new strategy was developed in which a global transformation function is estimated by only three control points with the best match. Therefore, the registration accuracy of the proposed method is much higher than that of traditional methods. The method involves the extraction of a set of CPs where misalignment between images can be expected to appear. These CPs are obtained by thresholding the gradient magnitude of the reference image and by applying the improved version of the well-known Harris corner detector. The corresponding set of CPs in the sensed image is determined by means of template matching, using the combined invariants-based similarity measure. Of all similarity measures developed so far, the combined invariants-based measure was demonstrated to satisfy the three constraints that have been emphasized in the introduction of this paper and therefore to be the most adequate measure for registration of remotely sensed images. After establishing the correspondence between small templates in the two images (reference and sensed), the central points of the corresponding templates in the sensed image are used as the corresponding set of CPs. The parameters of the affine transformation are computed by using only three pairs of CPs with the best match (minimum distance in the Euclidean space). The final correction is accomplished by warping the sensed image with respect to the

reference image by using the thin-plate spline-based interpolation. The performance of the proposed registration algorithm has been demonstrated by registering two multitemporal SPOT images and two multitemporal SAR images taken in different dates from one scene. In the experiments, a registration accuracy of less than 0.3 pixels has been achieved at each individual CP and an RMSE of less than 0.2 pixels has been obtained for the ten CPs. Moreover, the proposed registration has a better performance than the manual one, with an improvement in the registration accuracy of more than two pixels. This automatic algorithm is really fast and yields accurate global and local registrations. The approach is also robust, since it overcomes the difficulties of CP correspondence caused by the problem of feature inconsistency. The proposed registration technique may be applied to other satellite data as well as to airborne scanner data. In future work, the algorithm will be tested on other types of datasets, and its performance will be compared to other registration techniques.

REFERENCES

- [1] X. Dai and S. Khorram, "The effects of image misregistration on the accuracy of remotely sensed change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1566–1577, Sep. 1998.
- [2] —, "A hierarchical methodology framework for multisource data fusion in vegetation classification," *Int. J. Remote Sens.*, vol. 19, no. 18, pp. 3697–3701, Dec. 1998.
- [3] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, pp. 325–376, Dec. 1992.
- [4] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Apr. 2003.
- [5] H. Li, B. S. Manjunath, and S. K. Mitra, "A contour-based approach to multisensor image registration," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 320–334, Mar. 1995.
- [6] Q. Zheng and R. Chellappa, "A computational vision approach to image registration," *IEEE Trans. Image Process.*, vol. 2, no. 7, pp. 311–326, Jul. 1993.
- [7] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.
- [8] W. M. Wells, III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Imag. Anal.*, vol. 1, pp. 35–51, 1996.
- [9] P. Thevenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *IEEE Trans. Image Process.*, vol. 9, no. 12, pp. 2083–2099, Dec. 2000.
- [10] J. Le Moigne, W. J. Campbell, and R. F. Crompt, "An automated parallel image registration technique based on the correlation of wavelet features," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1849–1864, Aug. 2002.
- [11] J. Le Moigne and I. Zavorin, "Use of wavelets for image registration," in *Proc. SPIE Aerosense 2000, Wavelet Applications VII*, Orlando, FL, Apr. 24–28, 2000.
- [12] K. Johnson, A. Cole-Rhodes, I. Zavorin, and J. Le Moigne, "Mutual information as a similarity measure for remote sensing image registration," in *Proc. SPIE Aerosense 2001, Geo-Spatial Image and Data Exploitation II*, vol. 4383, Orlando, FL, Apr. 2001, pp. 51–61.
- [13] A. A. Cole-Rhodes, K. L. Johnson, J. Le Moigne, and I. Zavorin, "Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient," *IEEE Trans. Image Process.*, vol. 12, pp. 1495–1511, Dec. 2003.
- [14] J. Flusser and T. Suk, "A moment-based approach to registration of images with affine geometric distortion," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 382–387, Mar. 1994.
- [15] A. Goshtasby, "Registration of images with geometric distortions," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 1, pp. 60–64, Jan. 1988.
- [16] J. Ton and A. K. Jain, "Registering landsat images by point matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 27, no. 5, pp. 642–651, Sep. 1989.

- [17] B. S. Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration," *IEEE Trans. Image Process.*, vol. 5, no. 8, pp. 1266–1271, Aug. 1996.
- [18] J. P. Djamdji, A. Bijaoui, and R. Maniere, "Geometrical registration of images: The multiresolution approach," *Photogramm. Eng. Remote Sens.*, vol. 59, no. 5, pp. 645–653, May 1993.
- [19] Y. Bentoutou, N. Taleb, M. Chikr El Mezouar, M. Taleb, and L. Jetto, "An invariant approach for image registration in digital subtraction angiography," *Pattern Recognit.*, vol. 35, no. 12, pp. 2853–2865, Dec. 2002.
- [20] E. J. M. Rignot, R. Kowk, J. C. Curlander, and S. S. Pang, "Automated multisensor registration: Requirements and techniques," *Photogramm. Eng. Remote Sens.*, vol. 57, pp. 1029–1038, Aug. 1991.
- [21] X. Dai and S. Khorram, "A feature-based image registration algorithm using improved chain-code representation combined with invariant moments," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2351–2362, Sep. 1999.
- [22] P. Dare and I. Dowman, "An improved model for automatic feature-based registration of SAR and SPOT images," *ISPRS J. Photogramm. Remote Sens.*, vol. 56, no. 1, pp. 13–28, Jun. 2001.
- [23] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, Jun. 1989.
- [24] C. Schmid and R. Mohr, "Local gray value invariants for image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 5, pp. 530–535, May 1997.
- [25] J. F. A. Canny, "Computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
- [26] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *Int. J. Comput. Vis.*, vol. 37, no. 2, pp. 151–172, Jun. 2000.
- [27] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vision Conf.*, Manchester, U.K., 1988, pp. 147–151.
- [28] J. Flusser and T. Suk, "Degraded image analysis: An invariant approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 6, pp. 590–603, Jun. 1998.
- [29] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inf. Theory*, vol. 8, pp. 179–187, 1962.
- [30] W. H. Wong, W. C. Siu, and K. M. Lam, "Generation of moment invariants and their uses for character recognition," *Pattern Recognit. Lett.*, vol. 16, no. 2, pp. 115–123, Feb. 1995.
- [31] Q. Tian and M. N. Huhns, "Algorithms for subpixel registration," *Comput. Vis. Graph. Image Process.*, vol. 35, no. 2, pp. 220–223, Aug. 1986.
- [32] L. Van Tran and J. Sklansky, "Flexible mask subtraction for digital angiography," *IEEE Trans. Med. Imag.*, vol. 11, no. 3, pp. 407–415, Sep. 1992.
- [33] G. S. Cox and G. de Jager, "Automatic registration of temporal image pairs for digital subtraction angiography in medical imaging," in *Proc. SPIE Image Processing*, vol. 2167, Bellingham, WA, Feb. 1994, pp. 188–199.



Youcef Bentoutou received the Eng., Magister, and Doctorate degrees in electrical engineering from the University of Djilali Liabes, Sidi Bel Abbas, Algeria, in 1997, 2000, and 2004 respectively.

Since March 2002, he has been working as a research team leader at the National Center for Space Techniques, Arzew, Algeria. He is also an Associate Researcher in the Communication Networks, Architecture, and Multimedia Laboratory, University of Sidi Bel Abbas. From November 2000 to March 2002, he has been in training at the Surrey Space

Centre, University of Surrey. His principal research interests are in the fields of digital signal and image processing, medical and satellite imaging, and onboard data handling. His current work includes the development of image registration methods and the implementation of these methods on field-programmable gate arrays for application to satellite onboard data processing.



Nasreddine Taleb received the M.Sc. degree in computer engineering from Boston University, Boston, MA, the E.E. degree from Northeastern University, Boston, and Doctorat d'Etat degree in electrical engineering from the University of Sidi Bel Abbas, Sidi Bel Abbas, Algeria, in 1985, 1989, and 1999, respectively.

He is currently a Professor in the Department of Electronic Engineering, University of Sidi Bel Abbas, where he has been teaching since 1990. He is also a Senior Research Scientist in the Communication Networks, Architecture, and Multimedia Laboratory at the University of Sidi Bel Abbas. His principal research interests are in the fields of digital signal and image processing, medical imaging, and advanced architectures for implementation of DSP/DIP applications.

Dr. Taleb is a member of the IEEE Signal Processing Society.



Kidiyo Kpalma received the Ph.D. degree in image processing from the National Institute for Applied Sciences of Rennes (INSA), Rennes, France.

He is currently an Assistant Professor in the Department of Electronic and Computer Engineering at INSA. As a member of the Image and Remote Sensing group of the Institute of Electronics and Telecommunications of Rennes (IETR), a research unit linked to the French National Center for Scientific Research, his research interests are in the fields of image analysis, pattern recognition,

image fusion, and remote sensing.



Joseph Ronsin is currently a Professor in the Department of Electronic and Computer Engineering, Institut National des Sciences Appliquées de Rennes (INSA), Rennes, France. He is the Co-Responsible of the Image and Remote Sensing research group of the Institut d'Electronique et Télécommunications de Rennes, a research unit linked to the Centre National de la Recherche Scientifique. The INSA Image and Remote Sensing group focuses its research activities on representation and compression of video frames and sequences, analysis and interpretation of remote

sensing images, and prototyping on parallel and mixed architectures. He has also been responsible for several industrial grants from public and private laboratories.

XVII - Fusion de sur-segmentations

Ursani Ahsan Ahmad, Kidiyo Kpalma and Joseph Ronsin

Over-Segmentation of Feature Space for Initialisation of K-means Clustering

World Congress on Engineering and Computer Science 2007 - International Conference on Machine Learning and Data Analysis,

San Francisco (USA),

24-26 octobre 2007,

(7 pages)

Résumé

Nous proposons une approche originale pour améliorer l'initialisation des germes dans un processus de clustering par l'algorithme des k-means basée sur la méthode d'initialisation aléatoire de McQueen. L'approche introduite est testée à la fois, sur des données image ainsi que sur des données non-image. Pour les images nous exploitons l'énergie de Gabor pour extraire les attributs de textures afin de les appliquer à la segmentation d'images texturées. Des résultats encourageants ont été obtenus dans la segmentation d'une variété d'images de texture et de données non-image. Cette technique contribue ainsi à résolution du problème de la segmentation des données dans tous les domaines possibles, y compris le data mining, la reconnaissance de formes, l'intelligence artificielle ainsi que les statistiques.

Over-Segmentation of Feature Space for Initialisation of K-means Clustering

Ursani Ahsan Ahmad*, Kidiyo Kpalma, Joseph Ronsin

Abstract – This paper suggests over-segmenting the feature space for improved initialisation of the k-means clustering method. Two over-segmentations of a dataset are achieved and then fused to determine the best-possible initial means to perform the iterative process of k-means. The method is tested over both, image as well as non-image data. The image data is segmented using texture features based on Gabor energy features (GEF). The method is compared with single-run k-means and k-means with multiple restarts. Encouraging results have been obtained while segmenting a variety of texture images and the non-image data. The 40% over-segmentation is found to give best results. This technique addresses the issue of data clustering in all the possible areas including data mining, pattern recognition, machine learning and statistics.

Index Terms -- Fusion, K-means Clustering, Over-segmentation, Data mining

I. INTRODUCTION

Despite continued research by hundreds of researchers engaged for many long years, data clustering remains one of the most non-trivial and one of the most challenging tasks in the fields of data mining, statistics, bioinformatics, and image segmentation alike. Although there are many clustering techniques, however, so far k-means, due to its simplicity and faster convergence, remains the most popular among unsupervised techniques of data clustering. Nevertheless, often falling into a local minimum remains a problem *semper instans*. The k-means is generally performed several times (usually hundreds of times) before choosing the best solution among the several ones found therein. Moreover one is never sure that the best solution found is the optimal one.

This paper presents a method of initialising k-means clustering technique that helps prevent k-means from falling in the suboptimal points that lead to extremely undesirable clustering; thereby achieving optimal or near-optimal solution. A good comparison of the initialisation methods for k-means is [1] that discusses the problems associated with the k-means clustering.

The remainder of the paper is organized as follows: section II describes the 3 methods used and compared herein, section III describes more precisely the problems

and conditions of failure of k-means and the success of FOOS, section IV estimates the computational cost involved in the clustering methods, section V presents the results and their comparison and finally it section VI concludes with final remarks on the results and possible future work.

II. METHODS

This section presents the clustering results with single-run k-means and FOOS algorithms implemented with our own program and the *Gene Cluster 3.0*.

A. K-Means Algorithm

K-means clustering, invariant to instance order, based on the similarity metric of Euclidean distance was performed with the random initialisation. The dataset was randomly sampled to a much smaller subset, and then K means were selected randomly, making sure that there is a (considerable) minimum allowable distance between any two of the randomly selected means.

The iterative process of clustering is then performed to determine new class means and assign a class to each data-point in the dataset on the basis of closest Euclidean distance until difference between the current and the previous class means is zero. Generally, it is known that k-means assumes that the desired clusters are populated in spherical Gaussian distributions; and that any deviation from this situation causes k-means algorithm to fall into a sub-optimal point.

B. Gene Cluster 3.0

Gene Cluster 3.0 is a computer program developed by M. de Hoon [2] based on the original Cluster program written by Mike Eisen of Berkley Lab. In addition to the results obtained with our program of single-run k-means, results obtained with *Gene Cluster 3.0* are also presented and compared with those from FOOS. *Gene Cluster 3.0* performs k-means several times before choosing the best solution out of the several available. The criterion for choosing the best solution is the minimum Intra-Cluster Distance [1] defined as

$$E = \sum_{k=1}^K \sum_{n=1}^{N_k} |\mu_k - x_n|, \quad (1)$$

*Corresponding author: A. A. Ursani is Assistant Professor at Mehran University of Engineering & Technology, Jamshoro, Sindh, Pakistan, but since September 2005, he is PhD student at INSA de Rennes, France. Email: aursani@ens.insa-rennes.fr

where K is the total number of clusters, N_k is cardinality of the k^{th} cluster, μ_k is the mean feature vector of the k^{th} cluster, and x_n is the feature vector of the n^{th} data member of a running cluster k .

C. The Proposed Approach

The proposed approach, referred to as “FOOS”, is a means of finding the initial means that, unlike k-means, leads to the optimal or a near-optimal solution even if the desired clusters do not lie in the spherical Gaussian distribution. FOOS performs the k-means procedure only 3 times: it initialises the 3rd run of k-means with the cluster means found from the 2 preceding runs of the algorithm.

The clustering algorithm suggested here finds out 2 over-segmentations (OS) and then performs a kind of fusion of the clusters found therein. The k-means algorithm is first used to segment a dataset into M and N clusters, respectively, where K is the desired number of clusters, $M > K$ and $N = M + 1$. The 2 OS are then fused to determine good initial means to perform simple k-means clustering explained in the preceding section. Hence those K pairs of clusters, one from each M -cluster and N -cluster OS, respectively, are selected that have largest number of common member data-points.

Least possible OS is suggested for the datasets with $K \leq 5$. In this case, $M = K + 1$. The procedure is described mathematically in (2) through (7).

$$\bigcup_{m=1}^M X_m = \bigcup_{n=1}^N Y_n = I, \quad (2)$$

where $M = K + 1$, $N = K + 2$, I is the set of all the data-points in the dataset, X_m is one of the M clusters found in first OS and Y_n is one of the N clusters found in the 2nd OS.

$$\forall m, n, \exists Z_{m,n} = X_m \cap Y_n, \quad (3)$$

where $Z_{m,n} \in Z$, $|Z| = M \times N$ and $\|\cdot\|$ represents the cardinality of a set.

$$T_{m,n} = \|Z_{m,n}\| \quad m=1, 2, \dots, M \text{ and } n=1, 2, \dots, N \quad (4)$$

$$\begin{aligned} \forall m, \exists W_m = (m, n) : \max_{n=1}^N (T_{m,n}) \wedge \\ (m, n) : \min_{n=1}^N (\max_{m=1}^M (T_{m,n})) \notin W \quad \|W\| = K \end{aligned} \quad (5)$$

$$\bar{\mu} = \frac{1}{T_{W_n}} \sum_{i=1}^{T_{W_n}} \bar{x}_{W_n,i} \quad \forall I \mid 1 \leq I \leq k \wedge I \in N, \quad (6)$$

where W_n is the n^{th} order pair in the set W that is the set of order pairs representing the K largest clusters of common data-points from M and N segmentations; $\bar{\mu}_n$ and T_{W_n} , respectively, are mean and cardinality of the n^{th} cluster, and $\bar{x}_{W_n,i}$ is the feature vector of the i^{th} data-point in the n^{th} cluster

Equation (5) explains that the smallest sets are rejected, selecting only the K sets with highest number of members. After finding K largest clusters among $Z_{m,n}$, the FOOS algorithm computes the corresponding K means from them to be considered as initial means to perform the following and decisive iterative process of k-means over the dataset as follows:

$$x_j \in X_n \leftrightarrow d_{j,n} = \min_{i=1}^k (d_{j,i}) \quad (7)$$

where X_n is the n^{th} cluster, $d_{j,n}$ is the Euclidian distance of x_j from the n^{th} cluster, and $d_{j,i}$ is the Euclidian distance of x_j from i^{th} cluster.

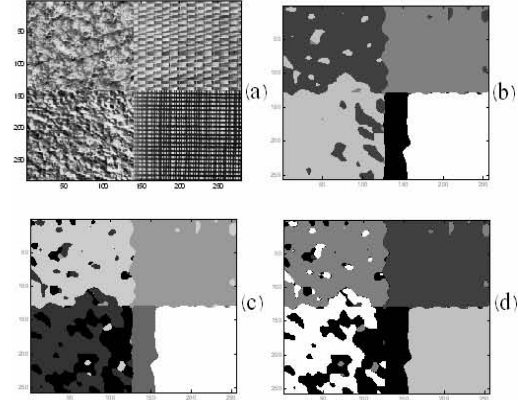


Fig.1. (a) The image with four textures from Brodatz collection, Segmentation into (b) 5 clusters (c) 6 clusters (d) The 4 largest classes of common member pixels

For example, fig. 1 shows an image with 4 (K) textures D92, D55, D4, and D21 from the Brodatz album, its segmentation into 5 (M), 6 (N), and 4 (K) clusters having the highest number of common member pixels. The pixels that are not part of the 4 largest classes of common members are shown in black. Table 1 shows the statistics of the number of common pixels in this particular case of the image in fig. 1. The table shows that class 3 from M segmentation and class 4 from N segmentation have the highest number of common member pixels, i.e. 16324. The following 3 (M, N) order pairs are (2, 5) with 13964, (5, 6) with 12777, and (4, 2) with 11135 common members. These sets of common members form the new clusters with the new means that are better means to be taken as initial means to perform k-means for clustering the dataset into K segments.

For the datasets having more classes, i.e. $K \gg 5$, the least possible OS doesn't seem to be the best solution, as shown in the results in section V. In case where $K \leq 5$, we performed least possible OS leading to $M = K + 1 = 6$, and $N = K + 2 = 7$; i.e. 2 more on 5 or 40% OS. If we extend this idea of 40% OS to a dataset with $K = 10$, the 2 proposed OS shall be $M = 13$ and $N = 14$.

Table 1: Statistics of the clusters formed after OS of the image in 5 clusters (left column) and 6 clusters (top row). The other cells in the table show the number of common pixels found in the 2 OS, with the highest 4 numbers in bold

Segmentation Into M (k+1) 5 clusters	Segmentation into N (k+2) 6 clusters								
	Class #	n		1	2	3	4	5	6
	m	$\ Y_n\ $		7395	11245	3656	16388	14075	12777
		$\ X_n\ $	$T_{m,n}$						
	1	3660		13	0	3647	0	0	0
	2	16710		2598	96	0	52	13964	0
	3	16405		38	10	0	16324	33	0
	4	15975		4746	11135	5	12	77	0
5	12786		0	4	4	0	1	12777	

III. NOTION BEHIND FOOS

Reference [3] is a recent work that points out 3 established problems/disadvantages with the k-means algorithm that are already pointed out by many earlier works. First that it requires the number of clusters beforehand, second that it is sensitive to the outliers, and third that any two randomly chosen initial means might be too close to be considered as two distinct centroids. Reference [1] notes that despite having all the advantages of convergence and computational simplicity, the k-means is affected by the choice of initial means and may easily converge into a local optimum due to its assumption that the clusters it is trying to find lie in a spherical Gaussian distribution.

Requiring the number of clusters beforehand is not always a problem. Often, the user prefers to decide upon it. Even the results of manual clustering of any given dataset will yield different outcomes since the number of clusters is almost always subjective depending on how rigorous segmentation is required. Therefore in many situations, the number of clusters (K) is a user-defined parameter. It is hence rather an advantage in many situations. The random selection of initial means can be repeated if any two of the selected means do not lie at a considerable distance. Rest of the problems, i.e. its assumption of spherical Gaussian distribution, sensitivity to outliers, and dependence on the initial means, all are inter-related. Keeping in view the determination of the number of clusters using Gaussian separation in [4], one can summarize these problems into a single problem statement as follows:

A given dataset usually has a different number of clusters from the Gaussian analysis [4] viewpoint than the number of clusters determined by application requires and/or subjective assessment. This disagreement makes distributions of the desired clusters look non-Gaussian and some of their members as outliers to the k-means algorithm, causing sensitivity to the choice of initial means while dividing the dataset into a desired number of non-Gaussian clusters.

The problem of so-called *outliers* is the same problem of non-Gaussian distribution in other words. For a given dataset, segmenting into a larger number of clusters relieves the problem of *outliers* or that of skewness in the cluster distributions; since more and smaller/finer Gaussian classes can more closely approximate non-Gaussian distributions. References [5-7] explain how non-Gaussian distributions can be approximated by piecewise Gaussian distributions.

Consider a situation requiring a dataset is desired to be segmented into 2 clusters, whereas the dataset seems to have 2 major clusters (large in population) "A", "B", and a minor (small in population) cluster "C" as per Gaussian separation [4]. The 2 major populations lie close in the feature space in such way that the minor cluster "C" is more distinct from cluster "A" than the major cluster "B" is from "A". In this situation, the k-means is most probably going to result in a merged class comprising the 2 major clusters "A" and "B" and the minor cluster "C" as a separate class. Even the k-means with multiple restarts will result in the same solution, since the merger of the 2 major classes will result in the minimum Intra-Cluster Distance.

Doing OS helps distinguish 2 relatively closer major classes from one another and the minor class. The following fusion process excludes the smaller/minor class, finding out the best initial means from the 2 major populations.

IV. COST ANALYSIS

The computational cost of k-means is given by

$$C_{kmeans} = O(KtN), \quad (8)$$

where K represents the number of clusters to be found, t is the number of iterations performed, and N is the cardinality of the data to be clustered.

Reference [8] argues that since $\{K, t\} \ll N$, the computational cost in (8) can be approximated by

$$C_{kmeans} = O(N). \quad (9)$$

Since FOOS can be thought of successive application of k-means to perform clustering into $K+2$, $K+1$, and K clusters respectively, its computational cost can be computed as

$$C_{FOOS} = O((K + (K + 1) + (K + 2))tN) \quad (10)$$

and considering that $3Kt \ll N$, can be approximated as

$$C_{FOOS} \cong O(3KtN) \approx O(N), \quad (11)$$

which tells that the computational complexity of the two methods happens to be the same.

The conventional k-means with multiple restarts is obviously computationally more demanding than FOOS that takes only 3 runs, whereas in the case of k-means, it is a function of cardinality of the dataset.

V. RESULTS

The results on both, the image and the non-image data are presented from the k-means, the *Gene Cluster 3.0*, and the FOOS algorithms.

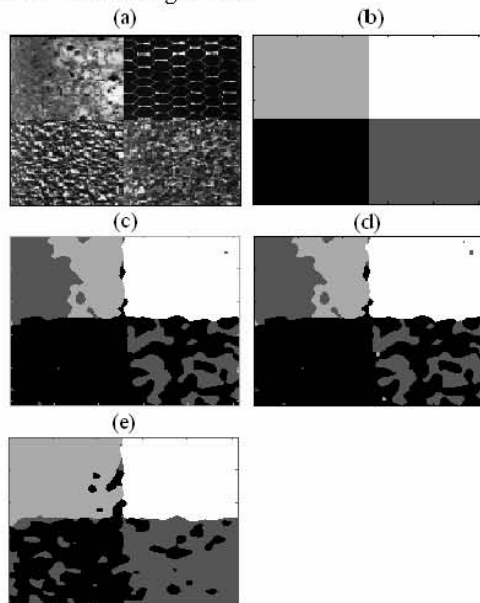


Fig. 2. (a) The image with 4 textures from Brodatz collection, (b) its ideal segmentation, segmentation result with (c) k-means, (d) Gene Cluster 3.0, (e) FOOS algorithm

A. Clustering the Image Data

For image data, the results are presented for the application of texture segmentation. The texture segmentation is one of the important aspects of not only the computer vision systems but also finds application in remote sensing (RS). The texture segmentation is carried out by extracting texture features and then using some data clustering technique such as k-means to segment the areas having homogeneous textures.

i. Collages of Four Brodatz Textures

An image with collage of 4 patches of the Brodatz textures D73, D34, D57, and D29, each of 128^2 pixels from Brodatz texture collection is shown in fig. 2(a). In this case, we over-segment the image into $M=5$ and then into $N=6$ clusters. The 2 OS are then fused to find 4 clusters. As shown in fig. 2(c), the k-means heavily misclassifies the textures on the bottom-right (D29) and the top-left (D73) corners achieving the accuracy of 69.5%. *Gene Cluster 3.0* is no different from k-means, achieving even worse, i.e. 69.2% accuracy. On the other hand, the FOOS algorithm considerably outperforms with an accuracy of 88.8% and is quite successful in discriminating all the classes.

In fact, many such combinations of textures were segmented. Table 2 lists the details of the collages and shows the results obtained with the k-means, FOOS, and the *Gene Cluster 3.0*. The *Gene Cluster 3.0* doesn't provide the optimal solution after making 100s of runs of k-means, and sometimes gives even worse results than k-means performed once by our program.

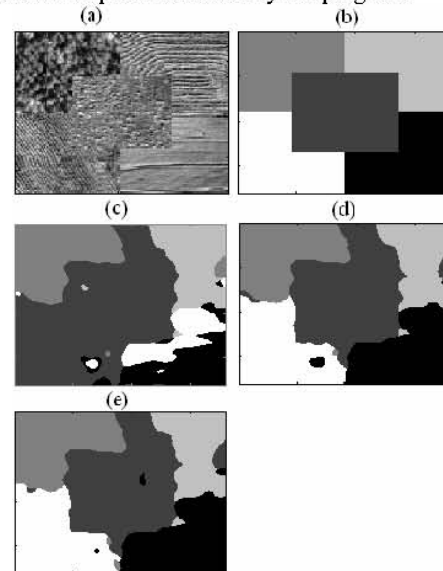


Fig. 3. (a) The image with 5 textures from airborne CASI sensors, (b) its ideal segmentation, segmentation result with (c) k-means, (d) Gene Cluster 3.0, (e) FOOS algorithm

ii. Five Textures From a Natural RS Scene

A collage of 5 textures from an RS image taken from an airborne sensor called CASI is presented. Fig. 3(c) shows that the k-means fails in discriminating between bottom-left texture and the central texture, causing a forceful creation of the 5th class from within the region of homogeneous texture on the bottom-right corner. Contrary to this, the FOOS segments all the textures successfully with most of the misclassifications in the top-right texture that has been confused with the central texture as shown in fig. 3(e). The result of *Gene Cluster 3.0*, shown in fig. 3(d) is slightly worse than

FOOS, achieving 89.1% accuracy after 26 runs of k-means. K-means achieves 65.8% accuracy as compared to 89.4% accuracy with FOOS.

In this example, k-means (single run) took 43 iterations, whereas FOOS took a total of 93 iterations that is far below 3 times those of k-means.

Table 2: The list and details of the images with a collage of 4 Brodatz textures each: The last column shows how many times the Gene Cluster found clusters using k-means before choosing the best among them

S. No.	Textures Selected (From top left to bottom right)	Segmentation Accuracy			
		K-Means	FOOS (3 runs of k-means)	Gene Cluster 3.0 Accuracy	No. of Runs
1	D73, D34, D57, D29	69.5%	88.8%	69.2%	274
2	D4, D55, D9, D21	67.6%	97.0%	96.9%	178
3	D24, D84, D4, D21	67.8%	95.4%	95.4%	453
4	D92, D55, D4, D21	67.5%	81.1%	67.2%	029
5	D3, D22, D112, D80	60.4%	91.5%	91.4%	192
6	D54, D84, D57, D100	52.3%	69.8%	69.5%	028
7	D73, D37, D57, D29	55.0%	73.2%	55.0%	050
Overall Average Accuracy		62.8%	85.3%	77.8%	172

iii. A Natural Image from Berkley Database

Here are the results on a natural image from Berkley Segmentation Database (BSD) reported in [9]. As shown in fig. 4(a), the image apparently has 3 textures, 1st one that of the grass, 2nd one that of the zebras, and 3rd in the rest of the image. The image has the size of 481X321. Fig. 4(b) illustrates its segmentation (into 36 segments) by a human subject (user #1107) as given in the BSD benchmark, clearly outlining the zebras, the grass, and rest of the image (the non-grass background). It should be noted however that the provided segmentation is with reference to the edge detections only, and not with respect to the texture. Fig. 4(c) and 4(d) show the result obtained with k-means and *Gene Cluster 3.0*, respectively, with the outline of fig. 4(b) overlaid on them. Fig. 4(e) shows the result obtained with FOOS. The *Gene Cluster 3.0* performed 472 k-means processes and took hours to complete.

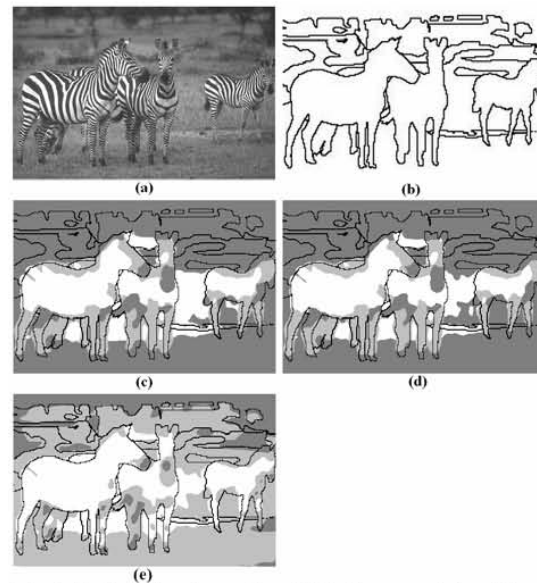


Fig.4. (a) The zebra image from BSD, its segmentation (b) as provided by BSD, (c) with k-means, (d) with *Gene Cluster 3.0*, (e) with FOOS

FOOS successfully separates the zebras, the grass, and the non-grass background. Having no texture, the mouths and the feet of the zebras also appear different from the zebra skin. Remaining misclassifications apparently indicate the reduced discrimination ability of the extracted features. On the other hand, the k-means and *Gene Cluster 3.0* fail in distinguishing between the grass and the non-grass background, forcefully creating a class within the zebra texture and classifying legs as different from zebra's main body. The k-means and *Gene Cluster 3.0* also misclassify the non-occluded part of the otherwise occluded zebra's back.

B. Clustering the Non-Image Data

Results are presented on 2 datasets from UCI machine learning repository [10].

As seen in the results on the image data, the single run k-means may result in very good or very bad clustering solution depending on its random initialisation. Hence, results from single-run k-means are considered no more.

i. A Dataset with 2 Classes

This is a database of human body dimensions, whose original source is [11]. The data include 507 observations of 21 body dimension as well as age, weight, height, and sex. The subjects include 247 male and 260 female individuals. The feature of sex was removed from this dataset. Further, all the feature values were rescaled from -250 to 250 to ensure that all the feature vectors get the equal weight in the distance measurement. The dataset was segmented into 2 clusters corresponding to the 2 sexes. The removed column of sex was used as the ground truth. The results from *Gene Cluster 3.0* and FOOS are exactly the same, i.e. 87.4%, with 443 instances classified correctly. Very small number of classes makes the task quite easy; therefore results ought to be the same. However, *Gene Cluster 3.0* performed 9216 runs of the k-means before reaching the solution, against 3 runs by FOOS.

ii. A Dataset with 10 Classes

This database was originally provided by E. Alpaydin and C. Kayna. The dataset contains 2000 instances corresponding to hand-written digits "0" to "9". The class distribution is uniform, viz. 200 instance of each class. The number of dimensions is 649. *Gene Cluster 3.0* stopped after performing just a single run of k-means, and as suggested, it was therefore run repeatedly 4 times. Each time, it stopped after performing a single run of k-means, yielding accuracies of 80.45%, 80.15%, 80.55%, and 80.60%, respectively. FOOS was performed 4 times with different values of M , as illustrated in table 3.

Table 3: FOOS Results with different over-segmentations

M	N	% Over-segmentation	Accuracy
11	12	20%	75.95%
12	13	30%	83.15%
13	14	40%	91.70%
14	15	50%	91.15%

The last 3 results, i.e. corresponding to 30-50% over-segmentations are better than those of *Gene Cluster 3.0*. The best result is the one corresponding to 40% over-segmentation, with an accuracy of 91.7%, i.e. 11% more than that of *Gene Cluster 3.0*.

Reference [12] reports that using neural network classifiers (NNC), even the learning error was not less than 10%, which corresponds to accuracies of under 90%! It further states that the combined NNC failed in distinguish between class "6" and class "9". The clusters

resulting from FOOS show that class "6" is well separated from class "9" with not a single member of class "6" classified as belonging to class "9" or vice versa; although both cluster "6" and cluster "9", have some misclassifications, with more than 90% accuracies.

VI. CONCLUSION

To initialize the iterative process of k-means clustering, an approach based on the fusion of over-segmentations is proposed. The approach has been shown to be successful with the segmentation of various image as well as non-image data that is publicly available. The favourable comparison is also shown with single-run k-means and *Gene Cluster 3.0*, reinforcing the theoretical notions behind FOOS presented in section III.

Since FOOS makes only 3 runs of simple k-means, its computational complexity is also much lower than the other proposed methods that require performing k-means exhaustively or repetitively on the multivariate data. Table 2 shows that *Gene Cluster 3.0* gives the same or worse results after ~172 runs as compared to FOOS that is equivalent to 3 runs of k-means.

To test its repeatability, the FOOS was performed many times repeatedly. It gave exactly the same results in some cases and only slightly different results in others. For example, it was always exactly the same result for the natural image of fig. 4.

Computationally, FOOS is found to take less iterations than 3 times those taken by a single run of k-means, ensuring that it is almost never more expensive than 3 runs of k-means.

Least possible OS was tried with the datasets containing up to 5 clusters. On the other hand, for a dataset with 10 classes, 40% OS was found to be the best rather.

ACKNOWLEDGMENT

The authors are thankful to Professor Kacem Chehdi and his research team TS12M at ENSSAT Lannion for granting permission to use the remote sensing images from airborne CASI sensors shown in fig. 3. We are also thankful to all the contributors in making available the datasets used herein.

Lastly, A. A. Ursani thanks Mehran University of Engineering & Technology, Jamshoro, Sindh, Pakistan for financial support for his research work at INSA.

REFERENCES

- [1] J. A. Lozano, J. M. Pena, P. Larranaga. An empirical, comparison of four initialization methods for the k-means, algorithm. *Pattern Recognition Letters* 20:1027-1040, 1999.
- [2] M. J. L. de Hoon, S. Imoto, J. Nolan, and S. Miyano. Open Source Clustering Software. *Bioinformatics*, 20 (9):1453-1454, 2004.
- [3] Bernard Chen, Phang C. Tai, R. Harrison, and Yi Pan. Novel Hybrid Hierarchical-K-means Clustering Method for Microarray Analysis, *Proceedings of IEEE Computation Systems Bioinformatics Conference Workshops*, 2005

- [4] Chang J.H., Fan K.C. and Chang Y. L. Multi-modal gray-level histogram modeling and decomposition. *Image and Vision Computing* (20), 203-216, 2002.
- [5] Alexander Steinwolf. Software for kurtosis and skewness fitting by way of piecewise-Gaussian approximation, *Bulletin of the International Statistical Institute, Contributed Papers of 49th ISI Session, Florence, Book 2*, p. 433-434, 1993.
- [6] Alexander Stienwolf. Approximation and Simulation of Probability distributions with a variable kurtosis value. *Computational Statistics and Data Analysis* (21):163-180, Elsevier, 1996
- [7] Alexander Stienwolf and Stephen A. Rizzi. Non-Gaussian Analysis of Turbulent Boundary Layer Fluctuating Pressure on Aircraft Skin Panels. *Jrnl. of Aircraft*, 43(6):1662-1675, 2006
- [8] Eduardo R. Hruschka, Estevam R. Hruschka, Thiago F. Co-voes, Nelson F. F. Ebecken. Feature Selection for Clustering Problems: a Hybrid Algorithm that iterates between k-means and a Bayesian filter. *Proc. the International Conference on Hybrid Intelligent Systems* (5):405 – 410, 2005
- [9] Martin and C. Fowlkes and D. Tal and J. Malik. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics, *Proc. 8th Int'l Conf. Computer Vision* (2):416-423, 2001 <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/BSDS300/html/dataset/images/gray/253027.html>
- [10] Murphy, P. M. and Aha, D. W. UCI Repository of machine learning databases [<http://mlearn.ics.uci.edu/MLSummary.html>] Irvine, CA: Univ. of Calif., Dept. of Information & Computer Sc., 1994.
- [11] Grete Heinz, Louis J. Peterson, Roger W. Johnson, and Carter J. Kerk. Exploring relationships in body dimensions. *Journal of Statistics Education*, 11(2), 2003
- [12] M. van Breukelen and R.P.W. Duin, Neural Network Initialization by Combined Classifiers. *Proc. 14th Int. Conf. Pattern Recognition* (1):215, 1998

XVIII - Descripteur de formes invariant par transformation affine

Mingqiang Yang, Kidiyo Kpalma and Joseph Ronsin

Affine Invariance Contour Descriptor Based on the Equal Area Normalization

IAENG International Journal of Applied Mathematics

May 2007

Vol. **36**, No 2, (6 pages)

Résumé

Dans cette étude, nous présentons un descripteur unidimensionnel de silhouettes d'objets à deux dimensions qui, en théorie, reste absolument invariant aux transformations affines. L'approche proposée se fonde sur l'aire enclavée dans un secteur fermé. Nous proposons alors une méthode de normalisation des contours. Après cette normalisation, le nombre de points sur un contour entre deux points désignés reste inchangé avec les transformations affines. Nous montrons que, pour un contour filtré, la surface d'un triangle dont les sommets sont le barycentre du contour et une paire de points successifs sur le contour normalisé à l'aide de cette approche est linéaire avec les transformations affines. Les résultats expérimentaux montrent que la méthode proposée est invariante : au choix du point initial sur le contour, aux transformations affines même dans le cas de fortes déformations et en présence du bruit. Nous proposons également une méthode pour simuler le bruit de contaminant requêtes avec un rapport signal sur bruit défini. En outre, la nouvelle méthode de normalisation peut être associée à d'autres algorithmes pour améliorer leur robustesse et pour diminuer leur complexité dans la mesure de similarité.

Affine Invariance Contour Descriptor Based on the Equal Area Normalization

Yang Mingqiang, Kpalma Kidiyo, Ronsin Joseph

Abstract— In this study, we present a one dimensional descriptor for the two dimensional object silhouettes which in theory remains absolutely invariant under affine transforms. The proposed descriptor operates on the affine enclosed area. We design a normalizing contour method. After this normalization, the number of points on a contour between two appointed positions doesn't change with affine transforms. We prove that for the filtered contour, the area of a triangle whose vertices are the centroid of the contour and a pair of successive points on the normalized contour is linear under affine transforms. Experimental results indicate that the proposed method is invariant to: boundary starting point variation, affine transforms even in the case of high deformations and noise on shapes in a given. We also propose a method to simulate the noise contaminating the test shapes and define the signal-to-noise ratio for a shape. In addition, the proposed normalization method can be associated to other algorithms for increasing their robustness to affine transforms and decreasing their complexity in similarity measurements.

Index Terms—Affine-invariant, Contour, Descriptor, Filter, Noise.

I. INTRODUCTION

The advent of multimedia and large image collections in different domains and applications bring a necessity for image retrieval systems. These systems are supposed to retrieve images in an effective manner based on a user's input query. Image retrieval is based on observation of an ordering of match scores obtained by searching through a database. The key challenges in building a retrieval system are the choice of attributes, their representations, query specification methods, match metrics and indexing strategies.

Considerable amount of information exists in two dimensional boundaries of objects which enable us to recognize objects without using further information. A shape is originally defined by x and y coordinates of its boundary points which are subject to changes if the camera is allowed to change its viewpoint with respect to the object, the resulting boundary of the object will be deformed. This deformation can be approximated by general affine transforms. As a result, a shape representation must be robust under similarity transformation which includes scaling, changes in orientation, shearing and translation. A number of shape representations have been suggested to recognize shapes even under affine transforms. Some of them are the extensions of well-known methods such as dyadic wavelet transform [1], Fourier descriptors [2], affine curvature scale space (CSS) [3], affine arc length [4] and moment invariants [5], etc. In these methods, the basic idea is to use a parameterization which is robust with respect to affine transforms.

Yang Mingqiang is with the IETR-INSA, UMR-CNRS 6164, Rennes, 35043, France. (Phone: +33 233 23 23 86 15; Fax: +33 233 23 23 82 62; E-mail: myang@ens.insa-rennes.fr)

Kpalma Kidiyo is with the IETR-INSA, Rennes, 35043, France. Ronsin Joseph is with the IETR-INSA, Rennes, 35043, France.

In this paper, to extract an affine invariant attribute, a new algorithm based on area equal normalization is proposed. In theory this attribute is absolutely robust under the affine transforms, even with serious transformations. Experimental results show that the proposed algorithm is also rather insensitive to noise. In addition, a high computationally efficient object recognition scheme is briefly presented.

II. FUNDAMENTAL CONCEPTS

A. Closed Curve

Let us consider the discrete parametric equation of a closed curve:

$$\Gamma(\mu) = (x(\mu), y(\mu)) \quad (2.1)$$

where $\mu \in \{0, \dots, N-1\}$. In general, curves of an application may have been parameterized with different number of vertices N .

B. Affine Transforms

The affine transformed version of a shape can be represented by the following equations [6]:

$$\begin{bmatrix} x_a(\mu) \\ y_a(\mu) \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x(\mu) \\ y(\mu) \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} + B \quad (2.2)$$

Where $x_a(\mu)$ and $y_a(\mu)$ represent the coordinates of the transformed shape. Translation is represented by matrix B , while scaling, rotation and shear are reflected in the matrix A . They are represented by the following matrices:

$$A_{\text{Scaling}} = \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix}, A_{\text{Rotation}} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, A_{\text{Shear}} = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

If S_x is equal to S_y , A_{Scaling} represents a uniform scaling. A shape is not deformed under rotation, uniform scaling and translation. However, non-uniform scaling and shear contribute to the shape deformation under general affine transforms.

C. Affine Invariant Parameters

There are two parameters which are linear under affine transforms. They are the *affine arc length*, and the *enclosed area*.

The first parameter can be derived based on the properties of determinants. It is defined as follows:

$$\tau = \int_a^b [x'(\mu)y''(\mu) - x''(\mu)y'(\mu)]^{1/3} d\mu \quad (2.3)$$

The second affine invariant parameter is enclosed area which is derived based on the property of affine transforms: Under an affine mapping, all areas are changed in the same ratio. Based on this property, Arbter et al. [2] defined a parameter σ , which is linear under a general affine transform, as follows:

$$\sigma = \frac{1}{2} \int_a^b [x(\mu)y'(\mu) - y(\mu)x'(\mu)] d\mu \quad (2.4)$$

where $x(\mu)$ and $y(\mu)$ are the coordinates of points on the

(Advance online publication: 24 May 2007)

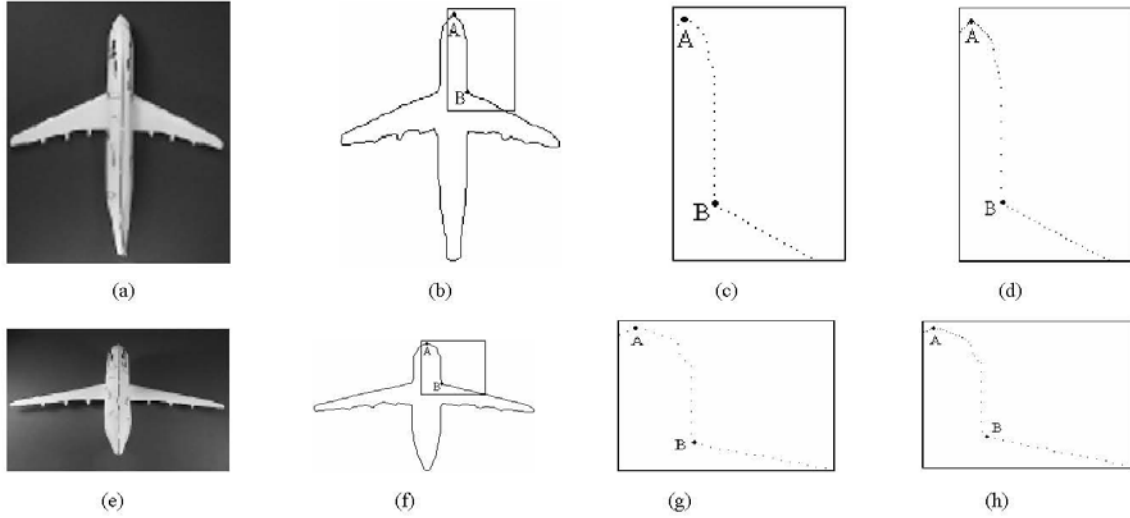


Fig.1. The comparison of equidistant vertices normalization and equal area normalization. (a) is the image of top viewpoint of a plane. (b) is the contour of image (a). (c) is a part of contour (b) normalized by equidistant vertices. (d) is a part of contour (b) normalized by equal area. (e) is the image of back top viewpoint of the plane. (f) is the contour of image (e). (g) is a part of contour (f) normalized by equidistant vertices. (h) is a part of contour (f) normalized by equal area.

contour with the origin of system located at the centroid of the contour. The parameter σ is essentially the cumulative sum of triangular areas produced from connecting the centroid to a pairs of successive vertices on the contour.

III. EQUAL AREA NORMALIZATION

All points on a shape contour could be expressed in terms of the parameter of index points along the contour curve from a specified starting point. With the affine transforms, the position of each point changes and it is possible that the number of points between two specified points changes too. So if we parameterize the contour using the equidistant vertices, the index point along the contour curve will change under affine transforms. For example, Fig.1(a) is the top viewpoint of a plane, and (e) is the back top viewpoint of the plane, so (e) is one of affine transforms of the image (a). Via region segmentation or edge following, we get the contours of the images (b) and (f). (c) and (g) are the parts of the contour (b) and (f) normalized by equal distance respectively. In Fig.1(c), the number of points between the appointed points A and B is 21; however, the number is 14 in Fig.1(g). So the contour normalized by equidistant vertices is variance to possible affine transforms.

In order to be invariance under affine transforms, a novel curve normalization approach is proposed that provides an affine invariant description of object curves at a low computational cost, while at the same time preserving all information on curve shapes. This approach is equal area normalization (EAN) and we present it as follow:

All points on a shape contour could be expressed in terms of two periodic functions $C(\mu') = (x(\mu'), y(\mu'))$, where

variable μ' is measured along the contour curve from a specified starting point.

1) Normalize the contour to N points with equidistance vertices. The new periodic functions are $C(\mu) = (x(\mu), y(\mu))$ and all the points on the contour are p_μ , where $\mu \in [0, N-1]$.

2) Calculate the second-order moments of the contour as its centroid G .

3) Transfer the contour to make centroid G to be at the origin of coordinates.

4) The last point $(x(N), y(N))$ is assumed to be the same as the first. Compute the area of the contour using the formula:

$$S = \frac{1}{2} \sum_{\mu=0}^{N-1} |x(\mu)y(\mu+1) - x(\mu+1)y(\mu)| \quad (3.1)$$

where $\frac{1}{2} |x(\mu)y(\mu+1) - x(\mu+1)y(\mu)|$ is the area of the triangle whose vertices are $p_\mu(x(\mu), y(\mu))$, $p_{\mu+1}(x(\mu+1), y(\mu+1))$ and centroid G (cf. Fig.2).

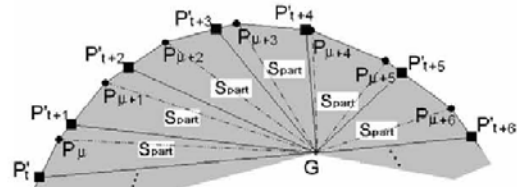


Fig.2. The method of normalization equal area. "•" is the vertex p of equidistant vertices normalization, and "■" is the point p' of equal area normalization. G is the centroid of the contour.

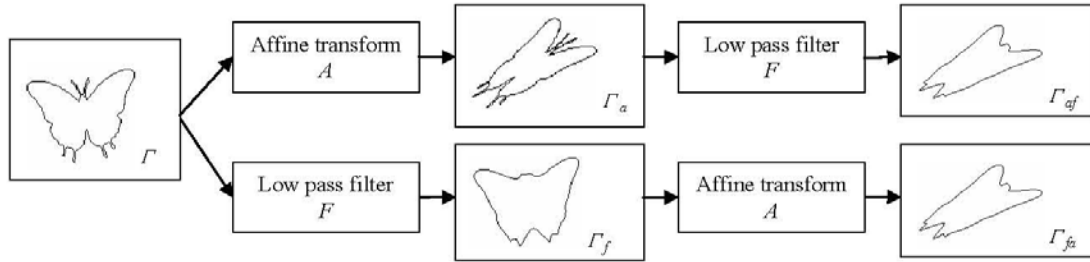


Fig.3. Demonstration of the theorem 1

5) Let the number of points on the contour after EAN be N too. Of course, any other number of points could be chosen. Therefore, after EAN, each enclosed area S_{part} defined by the two successive points on the contour and centroid equal to $S_{part} = S/N$.

Suppose all the points on the contour after EAN are p'_t . Let $C(t) = (x'(t), y'(t))$ represent the contour, where $t \in [0, N-1]$. Choose the starting point $p_0(x(0), y(0))$ on the equidistant vertices normalization as the starting point $p'_0(x'(0), y'(0))$ of the EAN, on the segment p_0p_1 , we seek for a point $p'_1(x'(1), y'(1))$, let the area $s(0)$ of the triangle whose vertices are $p'_0(x'(0), y'(0))$, $G(0)$ and $p'_1(x'(1), y'(1))$ equal to S_{part} . If there is not point to satisfy the condition, then we seek for the point p'_1 on the segment p_1p_2 . So the area $s(0)$, which is the sum of the areas of triangle p_0Gp_1 and triangle $p_1Gp'_1$, equal to S_{part} . If there is not the point to satisfy the condition yet, we continue to seek for this point at the next segment until the condition is satisfied. This point p'_1 is the second point on the normalized contour.

6) From point $p'_1(x'(1), y'(1))$, we use the same method to calculate all the points $p'_t(x'(t), y'(t))$, $t \in [0, N-1]$ along the contour. Because the area of each closed zone, for example, the polygon $p'_t[p_\mu p_{\mu+1} \dots] p'_{t+1} G$, $t \in [0, N-2]$ equal to S_{part} , the total area of $N-1$ polygon is $(N-1) \cdot S_{part}$. According to 5), we know it equal to $S - S_{part}$. So the area of the last zone $p_{N-1}' G p_0'$ equal to S_{part} exactly.

From Fig.2 we know, the area of the triangle $p'_t G p'_{t+1}$ equal to the area S_t of polygon $p'_t[p_\mu p_{\mu+1} \dots] p'_{t+1} G$ approximately if the distance between the two points $p_\mu p_{\mu+1}$ is close enough or the number N of the points on the contour is big enough. Therefore, we can use the points p'_t , $t \in [0, N-1]$ to replace the points p_μ , $\mu \in [0, N-1]$. Then the process of EAN is accomplished.

According to subsection II.C, after this normalization, the number of vertices between the two appointed points on a

contour is invariant under affine transforms. Fig.1(d) and (h) are the same parts of Fig.1(c) and (g) respectively. We notice that the distance between the consecutive points is not uniform. In Fig.1(d), the number of points between the appointed points A and B is 23, the number is also 23 in Fig.1(g). Therefore, after applying EAN, the index of the points on a contour can remain stability with their positions under affine transforms. This property will be very advantageous to extract the robust attributes of a contour and decrease complexity in the similarity measure. We can also use EAN to the others algorithm, to improve their robustness with affine transforms. For example, before do the algorithm of curvature scale space (CSS), the contour is normalized by EAN, all the maximum points in the image CSS will not change under affine transforms. This is beneficial to calculate the similarity between two CSS attributes.

In the following, we will study the part area S_{part} change with the effect of affine transforms and filtering.

IV. NORMALIZED PART AREA VECTOR

In this section, we look for the existing relations between the part area S_{part} , affine transforms and low-pass filtering.

THEOREM1:

If $\Gamma_a(x_a(\mu), y_a(\mu))$ is the transformed version of a planar curve $\Gamma(x(\mu), y(\mu))$ under an affine transform A , where μ is an arbitrary parameter, $\Gamma_{af}(x_{af}(\mu), y_{af}(\mu))$ notes that Γ_a is filtered by a linear low-pass filter F ; and if $\Gamma_f(x_f(\mu), y_f(\mu))$ notes that Γ is filtered by the same low-pass filter F , $\Gamma_{fa}(x_{fa}(\mu), y_{fa}(\mu))$ refers to the transformed version of Γ_f under the same affine transform A . Planar curve Γ_{af} is then the same as planar curve Γ_{fa} . In other words: $F[A(\Gamma)] = A[F(\Gamma)]$. (cf. Fig.3)

PROOF:

From Eq.2.2 we have

$$x_a(\mu) = ax(\mu) + by(\mu) + e \quad (4.1)$$

$$y_a(\mu) = cx(\mu) + dy(\mu) + f \quad (4.2)$$

Translation is represented by e and f . For all affine transforms contour, we transfer the centre gravity to origin. So the represents of translation e and f can be removed. Therefore, the affine transform can be represented by two simple formulae:

$$x_a(\mu) = ax(\mu) + by(\mu) \quad (4.3)$$

$$y_a(\mu) = cx(\mu) + dy(\mu) \quad (4.4)$$

The computation starts by convolving each coordinate of

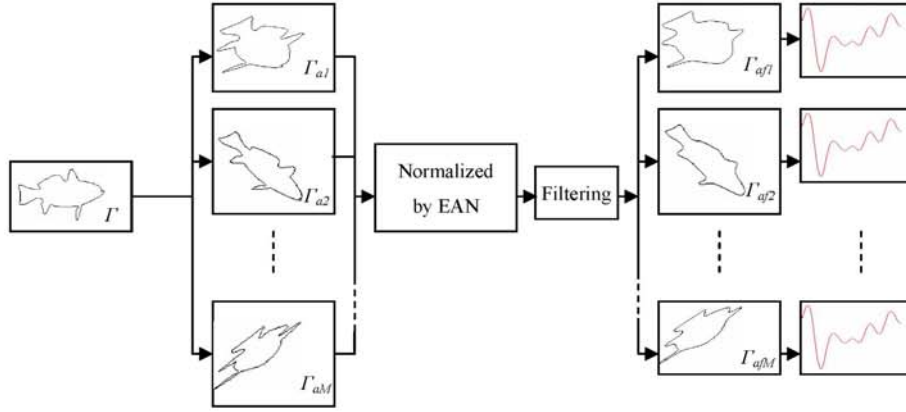


Fig.4. Demonstration of the theorem2

the curve $\Gamma_a(\mu)$ with a linear low-pass filter F whose impulse response is $g(\mu)$. In continuous form we have

$$\begin{aligned} x_{af}(\mu) &= x_a(\mu) * g(\mu) \\ &= [ax(\mu) + by(\mu)] * g(\mu) \\ &= ax(\mu) * g(\mu) + by(\mu) * g(\mu) \\ &= ax_f(\mu) + by_f(\mu) \end{aligned} \quad (4.5)$$

Where $*$ denotes convolution. Likewise,

$$y_{af}(\mu) = cx_f(\mu) + dy_f(\mu) \quad (4.6)$$

Compare Eqs.4.5, 4.6 and Eqs.4.3, 4.4, it is clear that point $(x_{af}(\mu), y_{af}(\mu))$ is also the point $(x_f(\mu), y_f(\mu))$ transformed by the affine transform A . So the planar curve Γ_{af} is same as the planar curve Γ_{fa} .

The theorem1 indicates that after exchanging computation order between affine transform and filtering, there is no change in the result.

THEOREM2:

For any affine transform of a closed contour, using EAN to set the parameters produces planar curve $\Gamma_a(x_a(t), y_a(t))$. If area $s_p(t)$ is the area of an enclosed sector whose vertices are a pair of successive points and the centroid of the contour and if $\Gamma_{af}(x_{af}(t), y_{af}(t))$ indicates that Γ_a is filtered by a low-pass filter F , then the changes in enclosed areas $s_p(t)$ on the Γ_{af} are linear with affine mapping. See Fig.4.

PROOF:

In section 3, we know the enclosed area $s_p(t)$ of the triangle on the contour filtered whose vertices are $(x_{af}(t), y_{af}(t))$, $(x_{af}(t+1), y_{af}(t+1))$ and the centroid G is

$$s_p(t) = \frac{1}{2} \text{abs}[x_{af}(t)y_{af}(t+1) - x_{af}(t+1)y_{af}(t)] \quad (4.7)$$

In the following, we will show that $s_p(t)$ is absolutely affine invariant to affine transforms. Due to the THEOREM1,

$$x_{af}(t) = ax_f(t) + by_f(t) \quad (4.8)$$

$$y_{af}(t) = cx_f(t) + dy_f(t) \quad (4.9)$$

and

$$x_{af}(t+1) = ax_f(t+1) + by_f(t+1) \quad (4.10)$$

$$y_{af}(t+1) = cx_f(t+1) + dy_f(t+1) \quad (4.11)$$

Therefore from Eq.4.7

$$s_p(t) = \frac{1}{2} \text{abs}\{ [ax_f(t) + by_f(t)][cx_f(t+1) + dy_f(t+1)]$$

$$\begin{aligned} &- [ax_f(t+1) + by_f(t+1)][cx_f(t) + dy_f(t)] \} \\ &= \frac{1}{2} \text{abs}[adcx_f(t)y_f(t+1) + bcx_f(t+1)y_f(t) \\ &- adx_f(t+1)y_f(t) - bcy_f(t+1)x_f(t)] \\ &= \frac{1}{2} \text{abs}(ad - bc) \cdot \text{abs}[x_f(t)y_f(t+1) - x_f(t+1)y_f(t)] \quad (4.12) \end{aligned}$$

Observing the Eq.4.12, $s_p(t)$ is just linearly proportional by a scale factor $\text{abs}(ad-bc)$. Accordingly, we have proved that enclosed areas $s_p(t)$ are linear with affine mapping.

DEDUCTION:

The proportion $v'(t)$ of the closed areas $s_p(t)$ with the total area S of the filtered contour is preserved under general affine transforms.

PROOF:

According to Eq.4.12, the total area S of the filtered contour is:

$$S = \frac{1}{2} \text{abs}(ad - bc) \cdot \sum_{t=0}^{N-1} \text{abs}[x_f(t)y_f(t+1) - x_f(t+1)y_f(t)] \quad (4.13)$$

So

$$\begin{aligned} v'(t) &= s_p(t) / S \\ &= \text{abs}[x_f(t)y_f(t+1) - x_f(t+1)y_f(t)] \\ &/ \sum_{t=0}^{N-1} \text{abs}[x_f(t)y_f(t+1) - x_f(t+1)y_f(t)] \quad (4.14) \end{aligned}$$

Observe the Eq.4.14, $v'(t)$ doesn't relate with the affine parameters a, b, c and d . Therefore $v'(t)$ is preserved under general affine transforms.

In addition, we can deduct major property of $v'(t)$: the integration of $v(t) = [v'(t) - 1/N]$ equal to zero.

PROOF:

$$\sum_{t=0}^{N-1} v(t) = \sum_{t=0}^{N-1} [v'(t) - \frac{1}{N}] = \sum_{t=0}^{N-1} \frac{s_p(t)}{S} - \sum_{t=0}^{N-1} \frac{1}{N} = \frac{S}{S} - \frac{N}{N} = 0$$

We refer to vector $v(t)$ as the normalized part area vector (NPAV). As theorem2 and its deduction show, in all cases, even those with severe deformations, the function $s_p(t)$ is also preserved, only the amplitude changes under general affine transforms; the NPAV $v(t)$ has an affine-invariant feature.

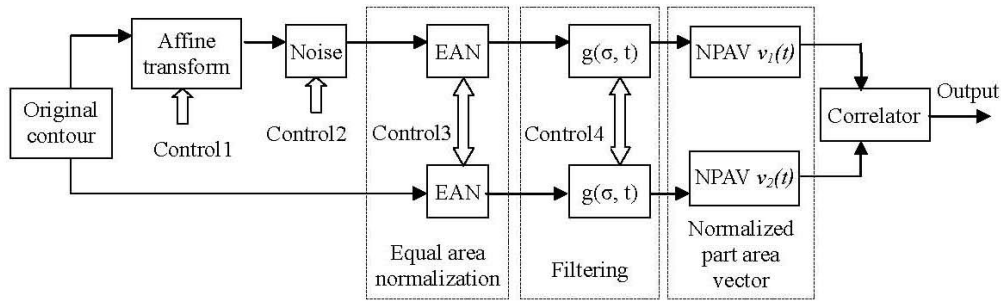


Fig. 5. The framework of experiments

V. EXPERIMENTAL RESULTS AND MULTISCALE NPAV

In experimental results, we will observe NPAV behavior in relation to affine transforms, EAN parameterization, filtering and noise. We experiment on the MPEG-7 CE-shape-1 database containing 1400 shape images. The framework of these experiments can be seen in Fig. 5. Here, the low-pass filter is a Gaussian filter with a scale parameter σ . The evaluation output is the maximum linear correlation coefficient between the NSAV of the upper pathway after affine transforms, noise, EAN parameterization and filtering and the NSAV of the lower pathway with the original contour.

First, we evaluate the effect of affine transforms on NSAV. Let the number of normalized points be 512, the position of the starting point be the same as the original contour and $\sigma=10$. According to subsection II.B, if the range of uniform scaling transforms and shearing transforms changes from 0.1 to 10 and the respective rotation angles are 60° , 120° , 180° , 240° and 300° , then all the linear correlation coefficients on NSAV have values above 97%.

We then assess the effect of equal area normalization (EAN) on NSAV. This experiment includes two aspects: on one hand the relation between NSAV and the number of points normalized by EAN and on the other hand the relation between NSAV and the position of the starting points on a contour. The results show that for a number of points on a contour normalized to 64, 128 and 256, and simultaneously with a starting point position displaced to 12.5%, 25%, 37.5% and 50% of the index points along the contour, all the linear correlation coefficient values stay above 97%.

Finally, we study the effect of noise on NSAV. When we resample the edges of an object, the original shapes are affected by noise and impairments, generating fluctuations on the boundary. The general method for noise simulation on a contour is to shift the x, y coordinates of the points on a contour independently. In this way, the order of certain points on the contour becomes confused (cf. Fig. 6(a)). However, in practice, such distortion cannot be present. In order to simulate the real effect of noise on an object, the displacement direction of each point affected by noise is controlled so that it is only vertical to the tangent of the point. Fig. 6(b) shows the effect of noise simulated with the proposed method i.e. there is a significant decrease in self-intersections.

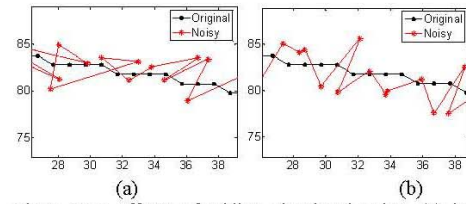


Fig. 6. Two effects of adding simulated noise. (a) is the effect of the general method of simulation. (b) is the effect of the proposed method

If its amplitude is controlled by a uniform random value, suppose the amplitude range is $[-r, r]$, and the average distance between all the points on the contour and its centroid is D , we define the signal-to-noise ratio (SNR) as follows:

$$\text{SNR} = 20 \lg \frac{2D}{r} \text{ (dB)}$$

We performed the experiments in such a way that the test contours were contaminated by random uniform noise ranging from high to low SNRs (cf. Fig. 7(a)-(d)).

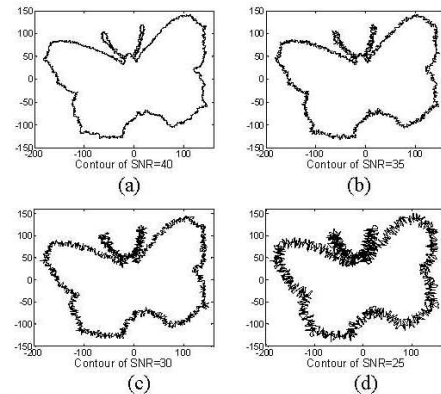


Fig. 7. Demonstration of a contour contaminated by various SNRs.

Table I. The average correlation coefficient of different SNR

SNR	Correlation coefficient
40dB	0.96366
35dB	0.96298
30dB	0.94907
25dB	0.89765

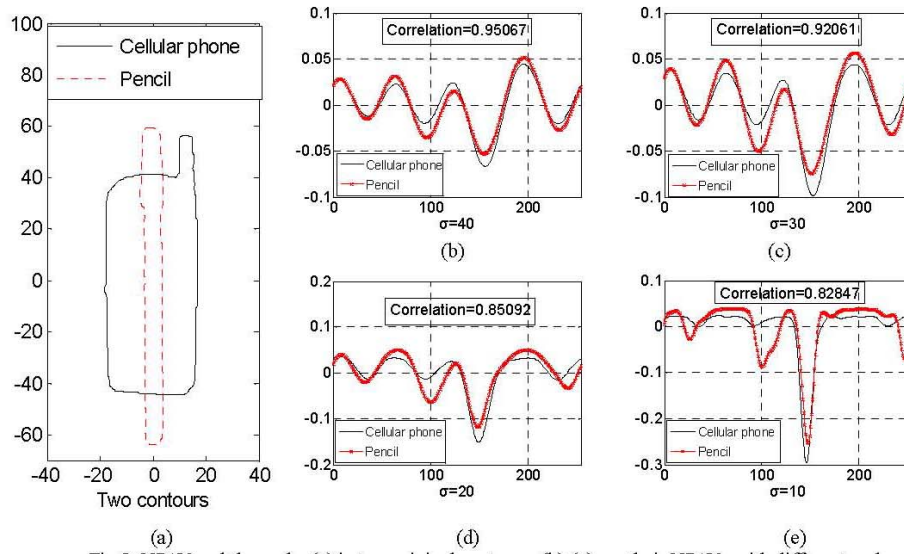


Fig.8. NPAV and the scale. (a) is two original contours. (b)-(e) are their NPAVs with different scale.

Table I shows the average correlation coefficient of all the NSAVs of shapes in the database with different *SVRs*. Correlation between NSAVs remains at an adequate level to allow efficient pattern recognition.

By analysing the results of experiments, we notice that NPAV is quite robust with respect to scale, orientation changes of objects, shearing, the position of starting point and the noise. Therefore, NPAV can be used to characterize a pattern for recognition purpose. In the following, we propose the object recognition scheme based on multiscale technology.

Under the high scale, it's possible that some of objects different have the similar NPAV. However the NPAVs of these objects have probably different feature under low scale. Fig.8 shows an example, the two objects in (a) are normalized by EAN to 256 points. Fig.8 (b)-(e) are their NPAVs under the scale $\sigma = 40$, $\sigma = 30$, $\sigma = 20$ and $\sigma = 10$ respectively. As can be seen in figures, the higher scale is, the more similar they are. On the contrary, the lower scale is, the more different they are. So we can present a shape with different scale: high-scale NPAV can be used to eliminate dissimilar shapes and low-scale images can discriminate between similar shapes.

In addition, as the NPAV expresses a contour with all the points normalized by EAN, it is rather redundant. To decrease the redundancy, we apply the wavelet algorithm to compress the NPAV. We have noticed that the higher the value of σ , the less high frequency remains in the NPAV, on the contrary, the lower the value of σ , the more high frequency remains in it. So we can use the high compress ratio to NPAV with higher σ , and low compress ratio to that with lower σ . In this way, a high computationally efficient object recognition scheme can be designed.

VI. CONCLUSION

In this paper, a new method of extracting invariants of an image under affine transform is described. Our representation is based on the association of the two

parameters the affine arc length and enclosed area, viz. we normalize a contour to invariant length by the affine enclosed area. We then prove two theorems and a deduction. They reveal that for a filtered contour, the part area is linear under affine transforms. We further define an affine-invariance vector: normalized part area vector (NPAV). A number of experiments on the MPEG-7 database demonstrate that NPAV is quite robust with respect to affine transforms and noise, even in the presence of severe noise. In addition, we propose a method to simulate the noise contaminating the test shapes and define the signal-to-noise ratio for a shape. Finally, a shape recognition scheme is briefly presented.

Future directions of the research include doing more extensive tests to assess retrieval results by multiscale NPAV, ameliorating the retrieval performance of the NPAN under part of the shape occluded.

REFERENCES

- [1] Erdem Bala, A. Enis Cetin, Computationally Efficient Wavelet Affine Invariant Functions for Shape Recognition, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.26, no.8, pp.1095-1099, 2004.
- [2] K. Arbter, W.E. Snyder, H. Burkhardt, G. Hirzinger, Applications of affine-invariant Fourier descriptors to recognition of 3-D objects, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.12, no.7, pp.640-646, 1990.
- [3] Mokhtarian, F. & Abbasi, S., Shape similarity retrieval under affine transform, *Pattern Recognition*, vol.35, no.1, pp.31-41, 2002.
- [4] D. Cyganski, R.F. Vaz, A linear signal decomposition approach to affine invariant contour identification, *Proceedings of SPIE-Intelligent Robots and Computer Vision X: Algorithms and Techniques*, vol.1607, pp.98-109, 1991.
- [5] A. Zhao, J. Chen, Affine curve moment invariants for shape recognition, *Pattern Recognition*, vol.30, no.6, pp.895-901, 1997.
- [6] Quang Minh Tieng, Wageeh W. Boles, Wavelet-Based Affine Invariant Representation: A Tool for Recognizing Planar Objects in 3D Space, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.19, no.8, pp.846-857, 1997.

XIX - Description de formes par analyse multiéchelle des points d'intersection

Kidiyo Kpalma and Joseph Ronsin

Multiscale contour description for pattern recognition

Pattern Recognition Letters,

October 2006

Vol.27, No.13, pp 1545-1559 (15 pages)

Résumé

Dans cet article, nous présentons une nouvelle méthode et les résultats préliminaires dans le contexte de l'analyse et la reconnaissance des formes. Basée sur une analyse multiéchelle d'une courbe, elle traite du contour des objets planaires. Elle utilise un filtre passe-bas gaussien pour lisser progressivement le contour en réduisant la largeur de bande du filtre. Après un ajustement d'échelle appliqué au contour lissé, il est remis à l'échelle du contour d'entrée : ainsi, les deux contours peuvent avoir des points d'intersection. En variant la largeur de bande du filtre et en observant les points d'intersection entre ces deux contours, nous produisons la Carte de Points d'Intersection (CPI) qui est caractéristique du contour.

Les résultats préliminaires obtenus en appliquant cette méthode à divers contours montrent que la carte CPI possède des propriétés très intéressantes dans le contexte de la reconnaissance des formes. Elle est insensible à translation et à la rotation et également résistante au changement d'échelle pour une gamme étendue. L'application des attributs extraits par cette méthode pour rechercher un modèle dans une base de données confirme l'efficacité de la méthode.



Multiscale contour description for pattern recognition

Kidiyo Kpalma ^{*}, Joseph Ronsin

IETR UMR CNRS 6164, Groupe Image et Télédétection, INSA de Rennes 20, avenue des Buttes de Coësmes, CS 14315, 35043 Rennes Cedex, France

Received 18 June 2004; received in revised form 26 January 2006

Available online 27 April 2006

Communicated by G. Borgefors

Abstract

In this paper, we present a new method and its preliminary results within the context of pattern analysis and recognition. This method is based on the multiscale analysis of a curve and deals with the contour of planar objects. Our method uses a low-pass Gaussian kernel to gradually smooth the contour by decreasing the filter bandwidth. Applying gain control to the smoothed contour stretches it to the same scale as the original one so that both contours intersect. By varying the bandwidth and marking all the intersection points between the smoothed contour and the original one we can generate the intersection points map (IPM) function.

The initial results obtained by applying this method to various contours appears to indicate that the IPM function has some very interesting properties within the context of pattern recognition. It is translation and rotation insensitive and also scale change resistant for a large range of scaling. The IPM function generated when applied to noisy contours shows that the method is resistant to noise for a range of noise energy.

Applying the features extracted by this method to retrieve a pattern from a database confirms the efficiency of the method.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Pattern analysis and recognition; Multiscale smoothing; Scale-space analysis; Curvature zero-crossing; Intersection point map

1. Introduction

Within the context of automatic pattern recognition, it is necessary to have a simple, fast and efficient algorithm especially when dealing with large databases. This algorithm must be resistant to rotation, translation, scale change and noise. Many techniques have been described in the literature related to this topic, however pattern feature extraction, which is the main problem in pattern recognition, is complex and an enormous amount of parameters are required to measure (Zhang and Lu, 2004; Sossa, 2000) shape area, elongation, moment invariants, number of holes, colour and texture,...

It is well known that objects, especially natural ones, consist of a more or less large range of scales; and that

the aspect of the object can change from one scale to another. Faced with this situation, it is very difficult to significantly describe a pattern using only one meaningful scale. To overcome this problem, increasingly more pattern analysis (or pattern description) techniques are based on multiscale or multiresolution representation methods (Lindeberg, 1994, 1998). Within this context, methods based on the pattern itself (Kadyrov and Petrou, 2001; Torres-Méndez et al., 2000; Grigorescu and Petkov, 2003; Belongie et al., 2002) exist as well as methods based on pattern contour behaviour (Matusiak and Daoudi, 1998; Wang et al., 1999; Roh and Kweon, 1998; Bruckstein et al., 1992, 1996; Mokhtarian et al., 1996; Mokhtarian and Mackworth, 1992; Latecki et al., 2000; Mokhtarian and Bober, 2003).

This study deals exclusively with pattern contour methods. In this paper, we present the initial results of a new method (MSGPR: a multiscale curve smoothing for generalised pattern recognition) (Kpalma and Ronsin, 2003)

^{*} Corresponding author. Tel.: +33 2 23 23 86 59; fax: +33 2 23 23 82 62.
E-mail addresses: kidiyo.kpalma@insa-rennes.fr (K. Kpalma), joseph.ronsin@insa-rennes.fr (J. Ronsin).

within the context of pattern analysis for pattern recognition. This scale-space (Matusiak and Daoudi, 1998; Wang et al., 1999; Bruckstein et al., 1992, 1996; Mokhtarian et al., 1996; Lindeberg, 1994; Mokhtarian and Mackworth, 1992) method is based on multiscale smoothing of a planar pattern contour. According to its framework, this method is totally translation insensitive. As we showed in the initial studies (Kpalma and Ronsin, 2003), the MSGPR method is also rotation insensitive and robust against scale change for a large range of scaling.

More sophisticated methods exist that deal with local invariant features, associated with planar curve smoothing: this enables to recognise occluded patterns by analysing the non-occluded part of their contour (Bruckstein et al., 1992; Trimeche et al., 2000). The main weakness of the MSGPR approach is that it fails to handle occluded contours and contours having undergone a non-rigid deformation. Nevertheless, when concerned by non-occluded planar contours, MSGPR method is advantageous in that it is simple and easy to apply.

The preliminary results indicate that the method is also resistant to additive noise.

By comparing pattern retrieval results obtained with this method with those provided by an adapted-CSS method, we are able to say that our method brings a positive point in the description of planar patterns.

After this introduction, in the next section, we recall the method structure. In Section 3, we present some of the main properties of the IPM function and we then go on to propose some interesting attributes that can be extracted from the IPM function for pattern characterisation. Section 4 concentrates on the application results within the context of pattern recognition by using the proposed attributes. Finally, in Section 5, we present our conclusions together with some discussions.

2. Description of the MSGPR method

The framework of this method can be broken down into four main stages as follows (see Fig. 1):

1. The input contour is separated into two parameterised functions.
2. Both functions are low-pass filtered (smoothed).
3. Gain control is then applied to both filtered functions so that the corresponding smoothed contour has the same scale as the input one.
4. Finally, the intersection points map (IPM) is constructed by detecting the intersection points of the input contour and the smoothed gain-controlled one.

2.1. Pre-processing

Before applying the MSGPR method, the input contour is normalised in relation to the contour length. This pre-processing method consists of the following three tasks:

- Contour resampling. The contour is resampled so that there are 360 equally spaced points. This is achieved by applying linear interpolation to the original contour points: if the number of points is greater than 360, this corresponds to sub-sampling. Using 360 equally spaced points produces points with a space of 1° between two consecutive points, in the circle limit case: of course, any other number of points can be chosen.
- Scale changing. Since the input contour points are pixels, their coordinates must be integer numbers. We therefore apply a scale factor to the resampled contour to generate a contour with 360 equally spaced points with approximately one unit distance between two con-

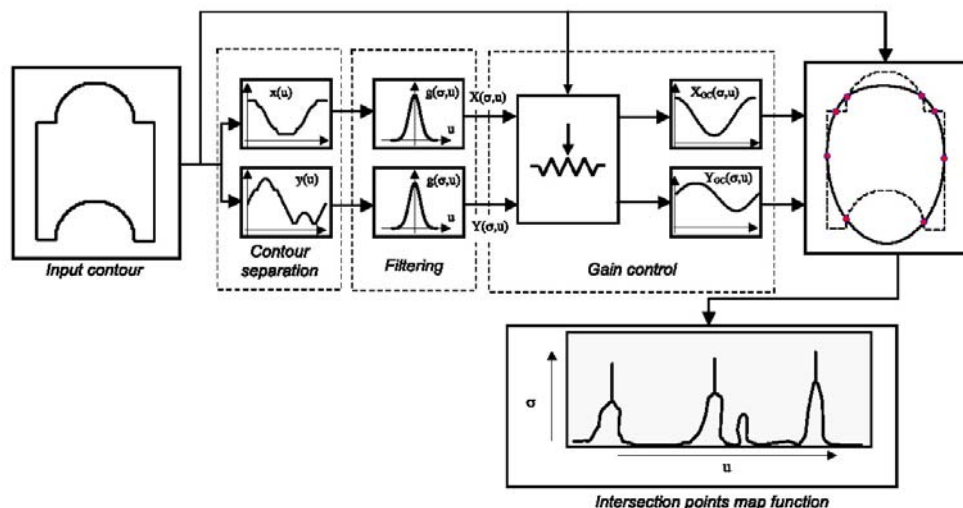


Fig. 1. MSGPR method diagram.

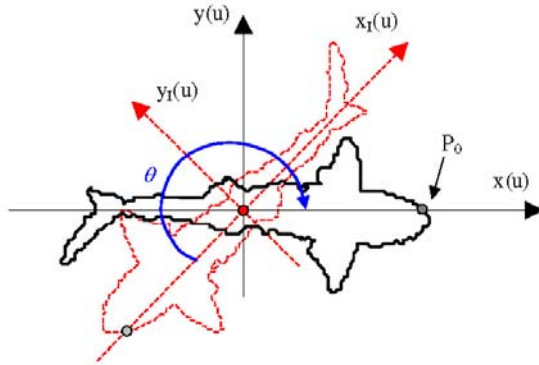


Fig. 2. Rotation compensation.

secutive points. The coordinates of these points are rounded up to integer values. After this pre-processing, the subsequent operations are real-valued so the results are no more rounded up to integer values.

- Rotation compensation so that the major principal axis is aligned with the x -axis and the point furthest from the centre of gravity is on the right: the centre of this rotation is set to the centre of gravity of the contour. This defines the P_0 starting point for subsequent processing (see Fig. 2).

NB: In the following sections, the expression “input contour” will be used to describe the pre-processed contour.

2.2. Coordinate separation

The input contour is represented by a series of points defined by their (x, y) coordinates. First, the input contour is separated into two functions $x(u)$ and $y(u)$ which are functions of the normalised curvilinear u parameter that varies from 0 to 2π relative to the curve length (see Fig. 3). Each point of the curve is then represented by its coordinates $(x(u), y(u))$.

For each contour, the origin of the u parameter is set to the P_0 starting point on the x -axis as defined previously. The points are then reorganised in a counter clockwise direction as illustrated in Fig. 3. Since the contour is a closed curve, the two functions are periodic with period $T = 2\pi$ relative to the u parameter.

2.3. Curve smoothing

Functions $x(u)$ and $y(u)$ are then gradually smoothed by decreasing the filter bandwidth. Similarly to the curvature scale space (CSS) method (Matusiak and Daoudi, 1998; Bruckstein et al., 1996; Mokhtarian and Mackworth, 1992) or other scale-space methods and according to the arguments stated in (Lindeberg, 1994), smoothing is based on the Gaussian filters $g(\sigma, u)$:

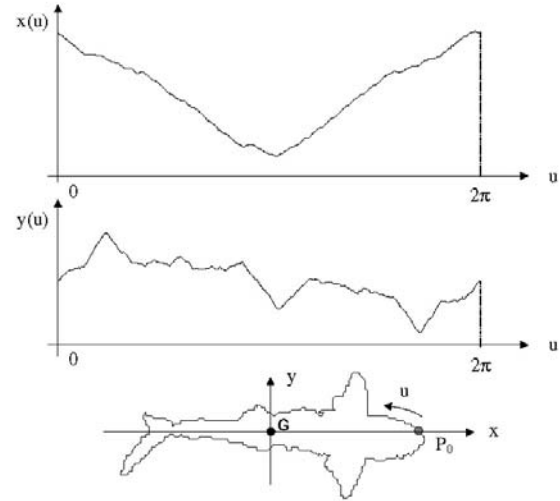


Fig. 3. A contour and its corresponding separated functions.

$$g(\sigma, u) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{u^2}{2\sigma^2}} \quad (1)$$

where σ is the standard deviation of the Gaussian kernel. The frequency response of the filter is then given by $G(\sigma, f)$ defined below

$$G(\sigma, f) = e^{-2\pi^2\sigma^2 f^2} \quad (2)$$

where f denotes the frequency variable. The -3 dB cut-off frequency of this filter is $f_c = \beta/\sigma$ where $\beta = \sqrt{\ln(2)}/(2\pi) = 0.1325$ and its bandwidth which equals $B = 2f_c = 2\beta/\sigma$ is inversely proportional to the standard deviation σ .

The filtered functions are then given by $X(\sigma, u) = g(\sigma, u) * x(u)$ and $Y(\sigma, u) = g(\sigma, u) * y(u)$ so that each $(x(u), y(u))$ point on the input contour leads to the $(X(\sigma, u), Y(\sigma, u))$ point on the output smoothed contour.

Since the bandwidth is conversely proportional to σ , it is clear that the bandwidth decreases as σ increases. Thus the filter cuts increasingly lower so that the output functions move towards their mean values when σ tends towards infinity. Fig. 4a shows the filtered functions for $\sigma = 10$ and $\sigma = 180$ with their corresponding smoothed contours. Due to this filtering process, the $X(\sigma, u)$ and $Y(\sigma, u)$ functions are attenuated so that the reconstructed smoothed contour does not have the same scale as the original one (see Fig. 4a).

2.4. Gain control

After low-pass filtering, the gain control system stretches the output contour so that it reaches the same scale as the input one and so that both curves intersect at certain points. The smoothed gain-controlled contour is then given by the coordinates $(X_{GC}(\sigma, u), Y_{GC}(\sigma, u))$ defined by

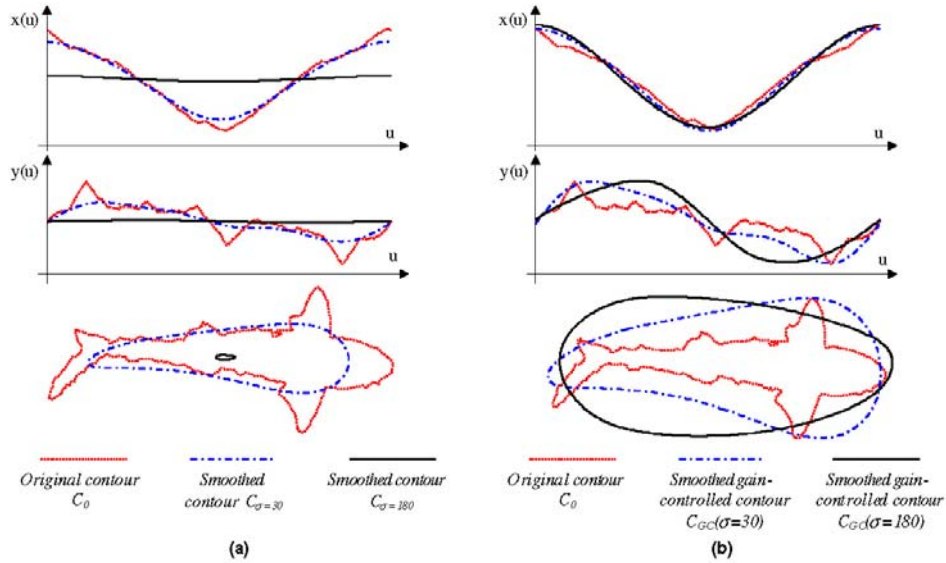


Fig. 4. Filtered functions and their corresponding smoothed contour.

$$X_{GC}(\sigma, u) = A_x[X(\sigma, u) - X_G(\sigma)] + X_G(\sigma) \quad (3)$$

$$Y_{GC}(\sigma, u) = A_y[Y(\sigma, u) - Y_G(\sigma)] + Y_G(\sigma) \quad (4)$$

where $(X_G(\sigma), Y_G(\sigma))$ are the coordinates of the centre of gravity of the smoothed contour. The $(A_x$ and $A_y)$ gain control coefficients can be determined in different ways and we can look out for the optimal one. For the results presented in this paper, A_x and A_y coefficients are defined as follows:

$$A_x = \frac{\sum_{n=1}^N |x(u_n) - x_G|}{\sum_{n=1}^N |X(\sigma, u_n) - X_G(\sigma)|} \quad \text{and} \quad A_y = \frac{\sum_{n=1}^N |y(u_n) - y_G|}{\sum_{n=1}^N |Y(\sigma, u_n) - Y_G(\sigma)|} \quad (5)$$

where $\{u_n\}$ are the N ($N = 360$) samples of the u parameter.

Fig. 4b shows the smoothed contour after the gain control system. The input and smoothed gain-controlled contours are now on the same scale so that they can intersect.

By increasing σ , the output gain-controlled contour moves towards a convex curve and has some intersection points with the input contour. By marking these intersection points for each σ , we obtain the intersection points map (IPM) function defined below which characterises the pattern.

2.5. Definition of the IPM function

After the gain control system, the IPM function is generated as follows. For each σ value, we define a function which is an image in the (u, σ) plane so that (see Fig. 5):

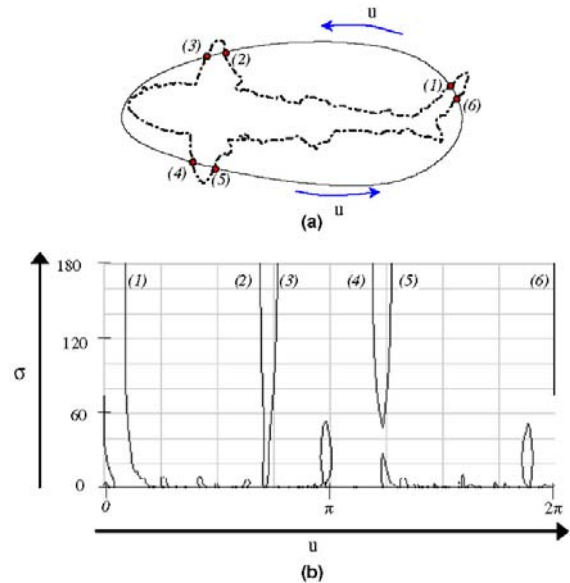


Fig. 5. Example of the IPM function.

- $IPM(u, \sigma) = 0$ (black) if the $(x(u), y(u))$ point is an intersection point between the original curve and the filtered gain-controlled one.
- $IPM(u, \sigma) = 1$ (white) if point $(x(u), y(u))$ is not an intersection point.

Fig. 5a shows the input and the smoothed gain-controlled one (that corresponds to $\sigma = 180$). This figure also

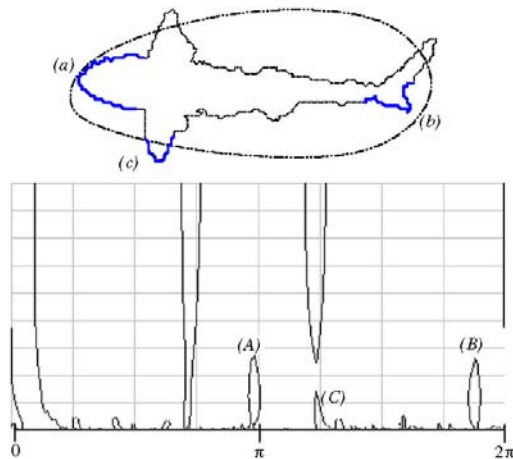


Fig. 6. Characterisation of convex/concave parts in the IPM function.

represents the IPM points for the corresponding σ . As σ increases, the number of IPM points decreases and we obtain “vertically” oriented lines or bands with variable widths. These remaining points constitute the points referred to as “limit IPM points”. In Fig. 5b, we represent the IPM function for σ , varying from 0 to 180, the IPM points indicated by (1)–(6) refer to the corresponding intersection points in Fig. 5a. Notice that for small σ values corresponding to large bandwidth filters, most of the contour points are IPM points.

By analysing Figs. 5 and 6, it appears that low values of σ produce fine details on a IPM function and high values produce coarse structures. In this way, we can say that high frequency information lies in the lower part of the IPM function since low frequency behaviour is found in its higher part. In other words, we can say that the IPM function contains all the information concerning the pattern from finer scales (details) to coarser scales. As σ increases, details are eliminated from the contour which becomes smoother and smoother so that the remaining intersection points can reasonably be expected to be significantly representative of the pattern being studied.

We also notice that the concave and convex parts of a contour are characteristic of the IPM function as illustrated in Fig. 6. In this figure, the convex parts labelled (a), (b) and (c) on the contour lead to the structures labelled (A), (B) and (C), respectively, on the IPM function.

3. Pattern characterisation

To show that this function is characteristic of the contour being studied, Fig. 7 presents two other patterns and their IPM functions. Fig. 7c represents the IPM function of the letter “M” (Fig. 7a) and Fig. 7d shows the IPM function of the letter “W” (Fig. 7b). As can be seen, these IPM functions are, somewhat, topologically similar. However, in terms of the distance between two consecutive IPM points at scale σ , we notice that these IPM functions are different from each another and also different from that

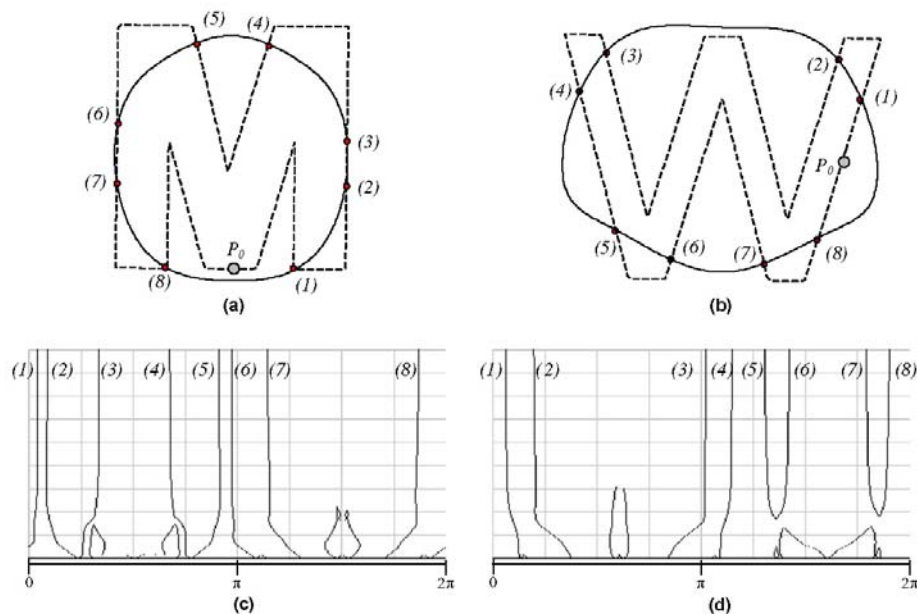


Fig. 7. Contours and IPM functions of letter “M” and “W” from the bold.chr character set from Borland®.

presented in Fig. 5. This shows that the IPM function is a good candidate to generate features for pattern recognition.

In the following sections, we will observe the behaviour of the IPM function regarding transformations such as scale change and rotation. We will then propose some features that can be derived from the IPM function.

3.1. The IPM function and scaling

In this section, we present the behaviour of the IPM function in relation of the scaling. The scaling transformation is obtained by applying a multiplicative coefficient to the contour coordinates. Fig. 8 shows that the IPM function is scale insensitive for a large range of scales. In this figure, the multiplicative coefficient next to the IPM function represents the scaling factor applied to the corresponding contour. According to the initial results obtained by applying the MSGPR method to the contour, the amount of limit IPM points remains constant and the IPM function does not change. For scale factors higher than the unit, the (entire) IPM function remains unchanged.

3.2. The IPM function and rotation

Fig. 9 shows a pattern (see Fig. 9a) and its rotated copies (see Fig. 9b–f). The rotation angles are 60° , 120° , 180° , 240° and 300° counter-clockwise. As can be seen in these figures, all IPM functions are topologically identical, except for a “circular” delay. In this way, rotating the contour around its centre of gravity is equivalent to a circular delay in the IPM function. By analysing the IPM functions obtained from a contour and its rotated copy, we can directly read the rotation angle between both patterns. This is illustrated

in Fig. 10. Since the u parameter is normalised to 2π , we can locate an IPM point A in one IPM function, determine its u_0 abscissa and then locate the corresponding IPM point A' in the second IPM function and also determine its abscissa u_1 . The γ rotation angle between both contours is then determined by the difference between both abscissa: $\gamma = u_1 - u_0$. Since the pre-processing method aligned the input contour with its major principal axis, the information related to the orientation is not available for subsequent processing. Therefore, to be able to directly read the rotation angle, we need to keep the compensated angle in mind during the pre-processing stage.

3.3. Feature definition and selection

With this IPM function, we can pre-determine a value for σ and use the number $N(\sigma)$ of IPM points as characterisation features. Since the number of IPM points is circular-delay insensitive, the two patterns in Fig. 7 are likely to be identical. To differentiate them, we need to use certain other features such as those suggested below.

By analysing the IPM functions obtained earlier, we notice that certain measurements remain invariant regardless of the transformation. These measurements can be used to characterise a pattern for recognition purposes. At present, we can suggest some heuristic features such as

- The number of IPM limits points (those corresponding to the highest value of σ);
- The set of distances between two consecutive IPM points at a given value of σ ;
- The minimum and maximum distances between two consecutive IPM limit points;

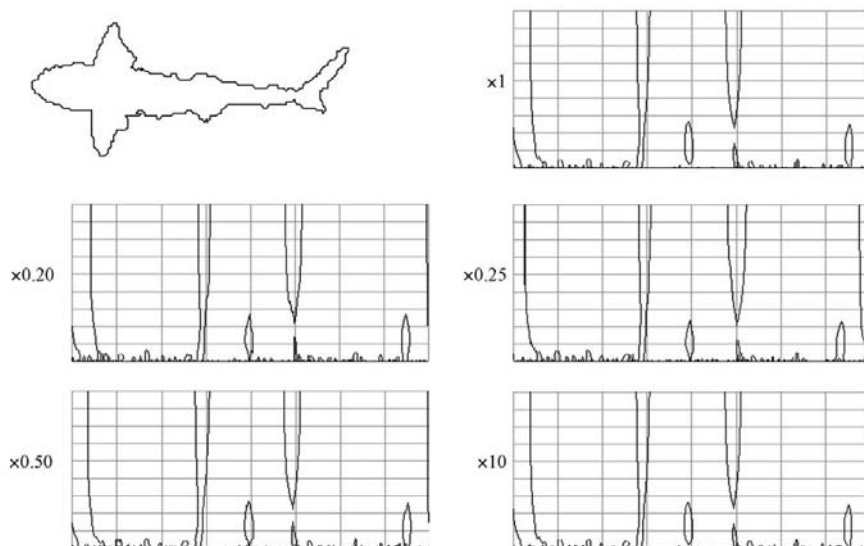


Fig. 8. Scaling effect.

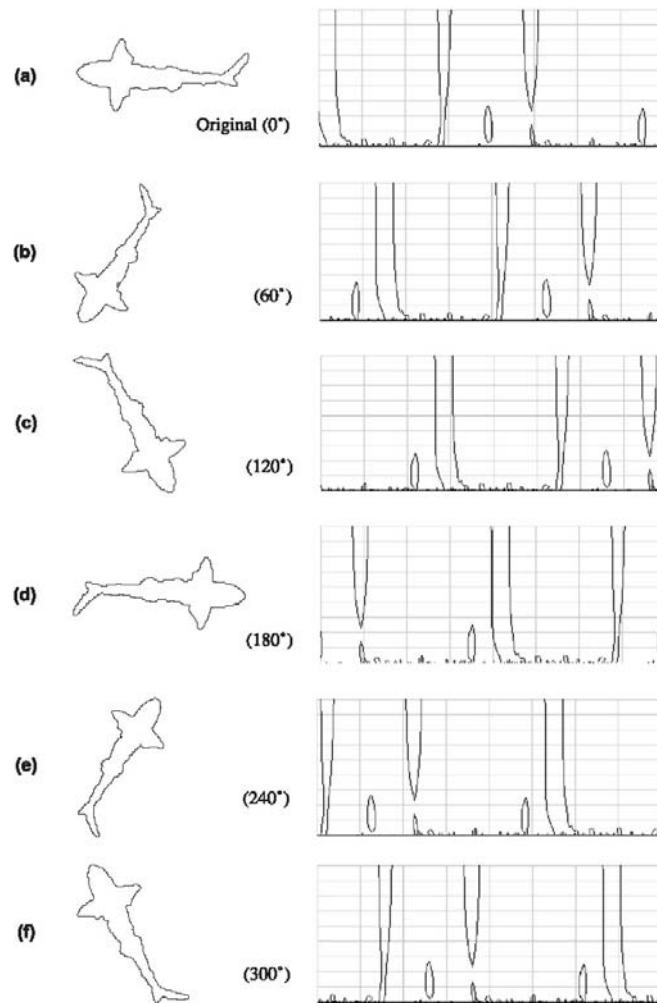


Fig. 9. Rotation effect.

- Measurements related to convex/concave parts of irregular objects. As mentioned earlier (see Fig. 6), the IPM function corresponding to irregular objects produces certain particular noticeable structures that can produce some interesting features such as their height, area and width at a fixed value of σ ;
- The ratio of arclength of arcs between two consecutive IPM points on the original contour over the Cartesian distance between both points for a given value of σ .

4. Application of the MSGPR method to pattern recognition

To evaluate the IPM function within the context of pattern recognition, we apply it to a pattern retrieval from a database. This database consists of 1167 different contours plus 100 transformed copies of the same contour (kk707)

and 25 transformed copies of 25 other contours. Transformations applied to the contours are affine transformations which are a combination of a rotation around the centre of gravity and scaling. Table 1 shows a set of 10 ($R_{i=0,1,\dots,9}$) rotation angles and a set of 10 ($H_{i=0,1,\dots,9}$) scaling factors. With these references, RmHnXYZ is obtained after applying an R_m rotation angle and an H_n scaling factor to the XYZ pattern. By way of example (see Table 1), the R7H3kk707 pattern is generated by applying a 270° rotation angle and a 0.75 scaling factor to the kk707 contour of the database. Fig. 11 provides a sample of the contours contained in our test database.

Note that during the transformation procedure, the output point coordinates are all integer values. This enables us to deal with points as if they were pixels produced by edge detection in a bitmap image. Of course, as we will see below, this is rather unfavourable to our approach.

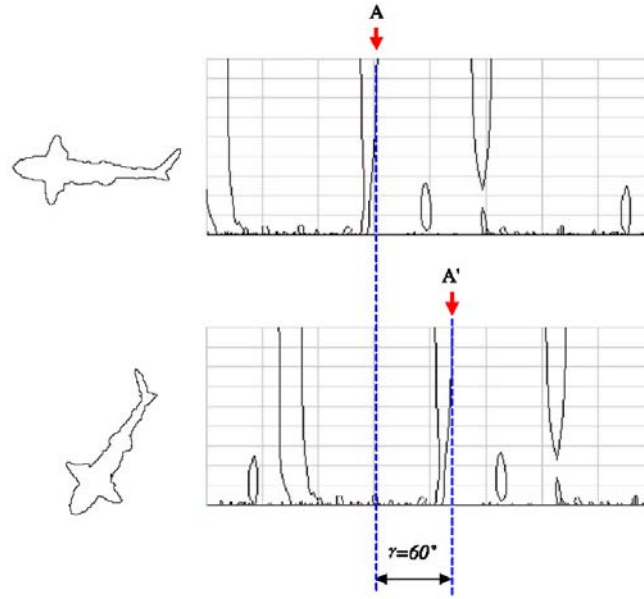


Fig. 10. Measuring the rotation angle.

Table 1
Rotation angles and scaling factors used in affine transformations

Rotation angle (R_i)	10°	45°	60°	90°	135°	180°	225°	270°	300°	330°
Scaling factor (H_i)	0.2	0.25	0.5	0.75	1	2	25	10	20	30
Index i	0	1	2	3	4	5	6	7	8	9

4.1. Attribute definition and selection

We can now describe and determine the features based on the IPM function that we used for our evaluation tests. To define these characterisation features, we first set σ to σ_0 value (e.g. $\sigma_0 = 180$). Then, for each pattern

- We consider the IPM points (at the set σ_0 value of σ) and select two consecutive p_a and p_b points which are, circularly, the furthest apart in the IPM function as illustrated in Fig. 12;
- We determine the distance between both points to produce the first d_1 component of the V_0 attribute vector;
- The next components of V_0 are distances coming after d_1 :

$$V_0 = (d_1, d_2, \dots, d_{N-1}, d_N) \quad (6)$$

Thus, the V_0 attribute vector is of dimension N where N is the number of IPM points at scale σ_0 . Using the most distant IPM points to define the attribute vector enables us to bypass the problem of correspondence between maxima of the interval tree as specified in (Mokhtarian et al., 1996).

To benefit from multiscale information of the IPM function, we can define a set of M values of σ ($\sigma_0, \sigma_1, \dots, \sigma_{M-1}$)

and determine the V_i attribute vectors ($i = 0, 1, 2, \dots, M-1$) corresponding to the σ_i scales (see Eq. (6)). The V global attribute vector is then produced by a concatenation of the individual V_i scale vectors as follows:

$$V = (V_0, V_1, \dots, V_{M-1}) \quad (7)$$

Each V_i vector is $N(\sigma_i)$ in length according to the number of IPM points on the σ_i scale. Thus, the length of the global attribute vector equals $N_T = \sum_{i=0}^{M-1} N(\sigma_i)$. Since V consists of vectors of multiple scales, it is considered to be a multiple scale attribute vector.

4.2. Similarity function

To measure the matching rate between two V_A and V_B attribute vectors associated to two different patterns, we define a similarity function as follows:

$$\text{SimScore}(V_A, V_B) = 50(1 + \cos(\gamma)) \frac{\text{Min}(\|V_A\|, \|V_B\|)}{\text{Max}(\|V_A\|, \|V_B\|)} \quad (8)$$

where γ is the angle between both vectors and where $\|\cdot\|$ indicates the module of a vector. This function ranges from 0% for very different vectors to 100% for perfectly matching vectors.

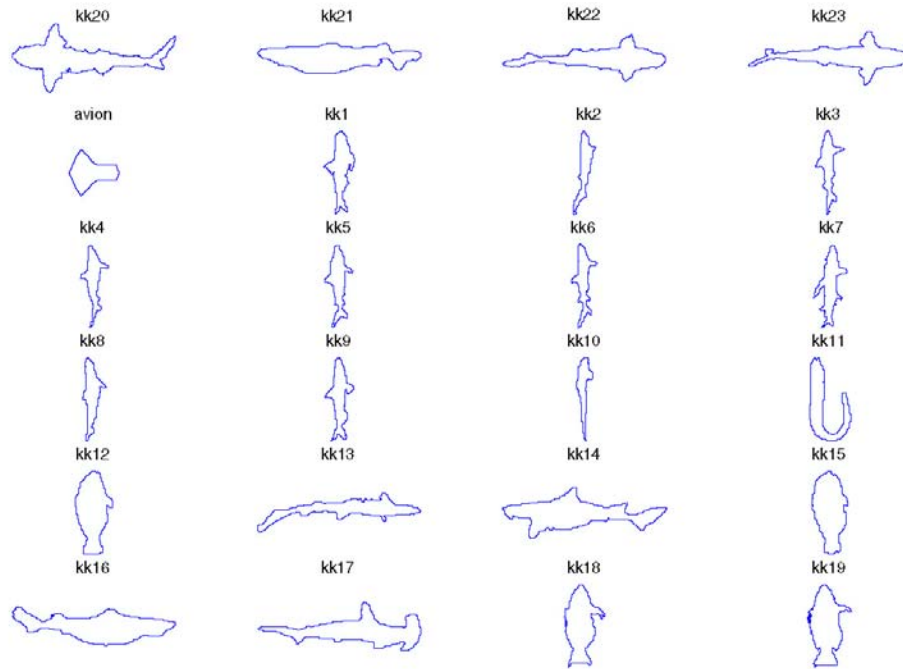
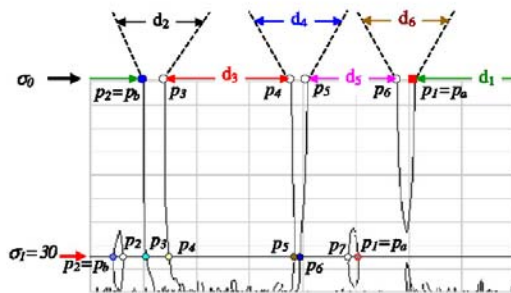


Fig. 11. An example of contours in the database.

Fig. 12. Attributes definition for a set of fixed σ values.

4.3. Results of pattern retrieval from a database

In this study, we used a set of five values for σ (180, 20, 30, 40, 50). Based on the total number of IPM points of the query contour that equals $N_T = 41$, in the case of a query from the kk707 family, 72 contours are retained for subsequent recognition stages. In this situation, the dimension of the V global attribute vector is equal to N_T .

After the pre-selection process based on the total number of IPM points, we apply the SimScore function to all pairs composed with the Q_0 query and to each of the selected contours (see Table 2). This pre-selection enables us to reduce computation time (at the SimScore function stage) by reducing the number of pairs to process.

Fig. 13 shows an example in which the Q_0 query pattern is a contour of the kk707 family i.e. a transformed copy of the kk707 pattern. This figure shows the eight most similar contours to the queried contour. The similarity scores for these matching contours are 100%. Table 2 shows the 72 selected contours and their similarity scores in relation to the queried contour.

As can be seen in Table 2 and Fig. 13, out of the 72 remaining contours, the first 32 have a similarity score of 100%. It is interesting to note that, similarly to the query contour, they all come from the kk707 family. Another interesting point to note is that the highest similarity score produced by a pattern from a different family is only 5%. The contour of the kk707 family with the lowest similarity score is the R6H1kk707 contour with a similarity score of 2%. This contour results from a (225°) rotation and a (0.25) scaling applied to the kk707 contour: the low similarity score is due to the scale reduction applied to the original contour. Fig. 14 shows the R6H1kk707 contour which is expanded to the original size and placed on top of the original rotated contour (with a rotation angle of 225°). This figure shows the defects introduced by the transformation applied to the original contour. These defects are due to scale reduction and mainly to the fact that the coordinates after transformation are rounded up to integer values (so that one point corresponds to one pixel in a bitmap image) as explained previously. This rounding up operation also introduces step-wise defects. These defects, which can produce peaks or self-intersections, are not dealt with in the

Table 2
Similarity scores for different contours in pattern retrieval from a database

kk707 (100)	R0H9kk707 (100)	R1H6kk707 (100)	R1H8kk707 (100)
R2H7kk707 (100)	R2H8kk707 (100)	R2H9kk707 (100)	R3H4kk707 (100)
R3H5kk707 (100)	R3H6kk707 (100)	R3H7kk707 (100)	R3H8kk707 (100)
R3H9kk707 (100)	R4H7kk707 (100)	R5H4kk707 (100)	R5H5kk707 (100)
R5H6kk707 (100)	R5H7kk707 (100)	R5H8kk707 (100)	R5H9kk707 (100)
R6H6kk707 (100)	R6H9kk707 (100)	R7H4kk707 (100)	R7H5kk707 (100)
R7H6kk707 (100)	R7H7kk707 (100)	R7H8kk707 (100)	R7H9kk707 (100)
R8H8kk707 (100)	R8H9kk707 (100)	R9H8kk707 (100)	R9H9kk707 (100)
R0H6kk707 (98)	R1H9kk707 (98)	R4H6kk707 (98)	R8H7kk707 (98)
R9H6kk707 (98)	R0H8kk707 (98)	R6H7kk707 (98)	R6H8kk707 (98)
R4H9kk707 (97)	R8H6kk707 (97)	R2H6kk707 (95)	R4H8kk707 (95)
R1H7kk707 (93)	R2H5kk707 (87)	R9H7kk707 (87)	R8H5kk707 (85)
R7H3kk707 (76)	R9H5kk707 (74)	R6H2kk707 (70)	R4H5kk707 (66)
R0H5kk707 (62)	R8H4kk707 (55)	R6H4kk707 (54)	R3H2kk707 (52)
R5H2kk707 (47)	R2H2kk707 (45)	R1H2kk707 (29)	R8H1kk707 (28)
R5H0kk707 (16)	kk1097 (5)	R8H0kk707 (4)	kk882 (4)
kk333 (3)	kk850 (2)	kk929 (2)	kk902 (2)
kk829 (2)	R6H1kk707 (2)	kk511 (1)	kk916 (1)

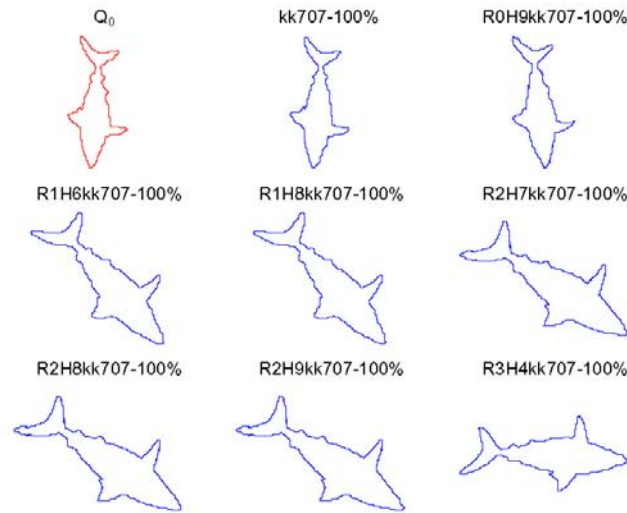


Fig. 13. kk707 recognition results for a set of five σ values.

present study. If these defects were taken into account, the results could be improved.

4.4. Comparison with an adaptation of CSS method

To evaluate our method, we have adapted the CSS (curvature scale space) method. The maxima of the CSS curves and their location are then used like shape descriptor. Only maxima between the major maximum M and the maximum $m = 0.2M$ are taken into account to define the attribute vector. We then apply the similarity function of Eq. (8) after a pre-selection based on the length of the attribute vector as it is explained in Section 4.3. To compare the retrieval results from both methods, we carried out three experiments:

- (1) The query shape “vase” is not in the database. The goal of this experiment is to assess the ability to discriminate shapes. This enables us to evaluate the sensitivity to false alarms situations in pattern recognition.
- (2) The query pattern “kk707” is in the database and there are also 100 members of the same family in the database. With this experiment, we aim at assessing the separability of the methods: how far are the less similar pattern from the query family and the most similar pattern from a different family (a non-parent pattern of the query).
- (3) We have applied both methods to a sample contours of the MPEG-Core Experiment CE-Shape-1. We applied them to retrieve a pattern from a database

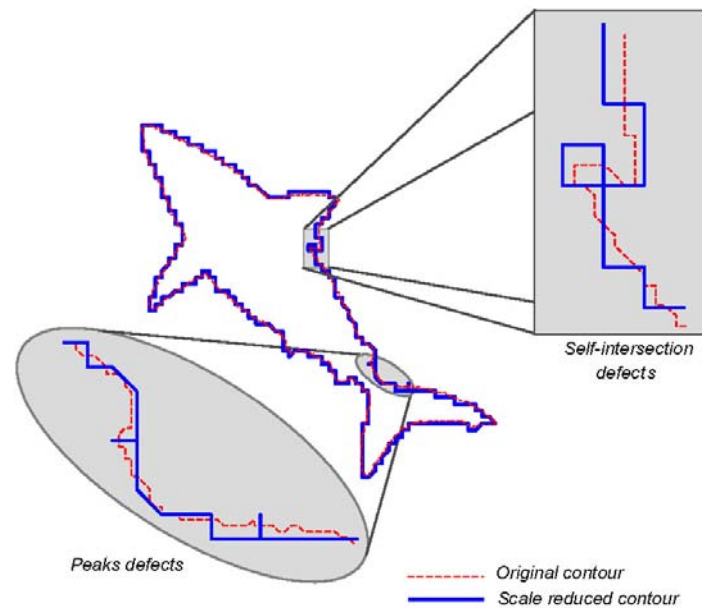


Fig. 14. Illustration of the defects introduced during transformation: step-wise defects, peak defects and self-intersection defects.

consisting of 20 contours of the same pattern that was subject of different kind of rigid or non-rigid transformations. Fig. 17 shows an example of such contours: in this example, contours are extracted from a flying bat in different kind of motion.

From these experiments, one can note the following interesting points:

- (1) When the query does not exist in the database (see Fig. 15), the IPM-based method retrieves 12 patterns with 26% maximal similarity score while the adapted CSS-based method retrieves 58 patterns with 50% maximal similarity score. Our method gives lower maximal similarity score (26%) than the CSS-based method (50%). In this experiment, the lower the maximal similarity score is and the better is the recognition system: indeed when the similarity score is low, one can decide easily that the query pattern is unknown (does not exist in the database).
- (2) In the second experiment (see Fig. 16), the lowest similarity score obtained from a parent of the query pattern is 91% and the highest score obtained from a non-parent pattern is 43%: the difference between both scores is $\Delta s_1 = 48\%$. With the adapted CSS method, this difference is $\Delta s_2 = 1\%$ while the lowest similarity score from a parent is 69%. Of course, a large difference indicates a good separation between classes: in other words, the best method is that method which provides large difference.

- (3) In the third experiment (see Fig. 18), when applied to contours of bats, the adapted CSS method yields better results than our method. As it can be seen in Fig. 17, some changes from a contour to another are not rigid transformations but they are caused by non-rigid deformations due to the motion of the animal. Since our method is essentially a global-driven method, it may fail for more or less non-rigid deformations. It is adapted better for rigid deformations as define for part A experiment in (Latecki et al., 2000; Mokhtarian and Bober, 2003). In this article, the authors have considered three different stages according to the three parts of the MPEG-7 Core Experiment CE-Shape-1: the part A that deals with robustness to scaling and to rotation is defined as a necessary condition to satisfy. In accordance with these experiments, we can reasonably say that our method is a good candidate to satisfy this necessary condition in shape description.

In any case, the query, itself, when it is present in the database, is retrieved with a score of 100%. As shown in Fig. 18, bat-16, which is the query, is retrieved with a score of 100%. Bat-17 is retrieved with a similarity score of 89%: this high similarity score indicates that both patterns are similar as it can be seen visually in Fig. 18. The next contour, bat-01, is found to be similar to the query with 40% similarity score. Visually, both contours are very similar but bat-01 has some high frequency details on the upper side of its wings: that explains the low score. Bat-05 and bat-04 present a very weak score (8% and 5%, respectively),

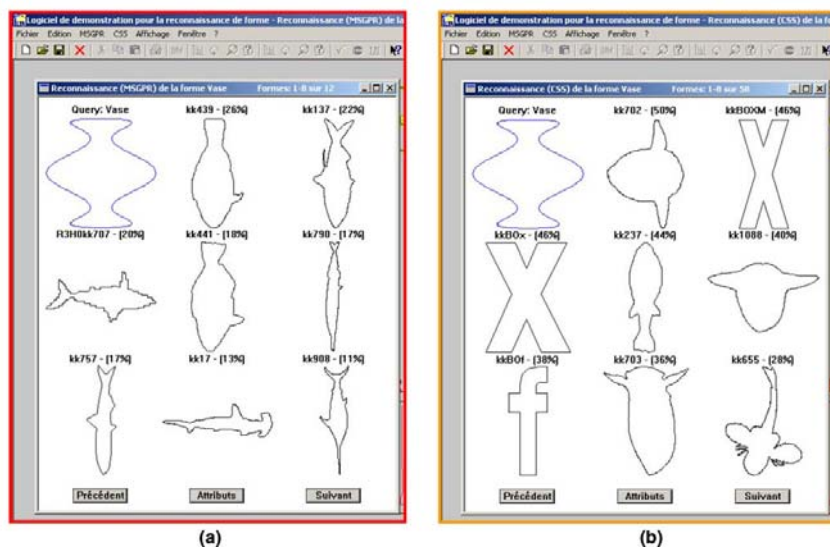


Fig. 15. Retrieval of a pattern that does not exist in the database: (a) results from IPM-based method and (b) results from the CSS-based method.

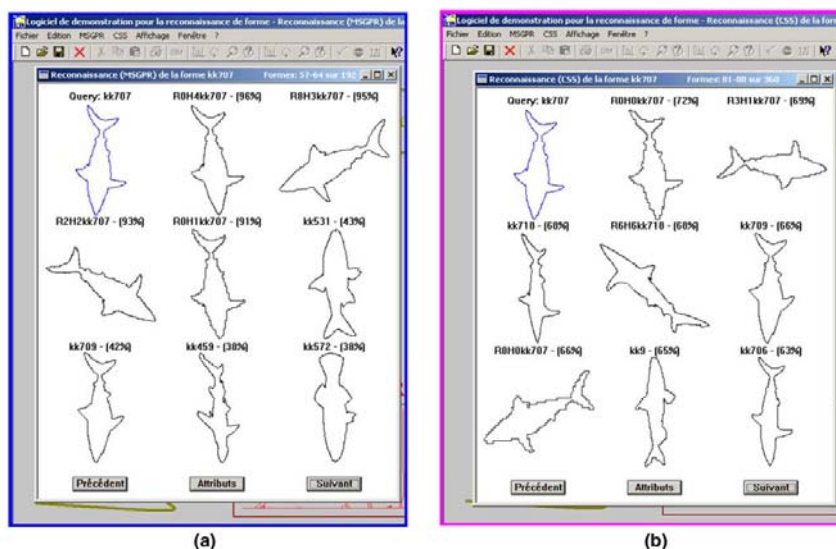


Fig. 16. Retrieval of a pattern from the database containing some derived patterns (by applying rotation and/or scaling) of the query: (a) results from IPM-based method and (b) results from the CSS-based method.

because contrarily, with bat-01, their ears are quite apparent: what creates this strong dissimilarity with the query.

Table 3 summarises results that enable us to carry out a comparison of both the adapted CSS-based method and the MSGPR method. In this study, we used three families of contours: bats, bell and children extracted from the MPEG-7 Core Experiment CE-Shape-1 database. In each family, there are 20 contours representing different views or deformations due to the object movements. For a given

family, we use each contour as a query and then record the minimal and maximal similarity score and the number of contours that are retained at the first step.

The average value of each of these three parameters is then registered in Table 3. For each method, the average value of the maximal similarity score is 100% indicating that the query contour is always recognised with 100% similarity score. This parameter shows the capability of a method to recognise a pattern when this one exists in the database.

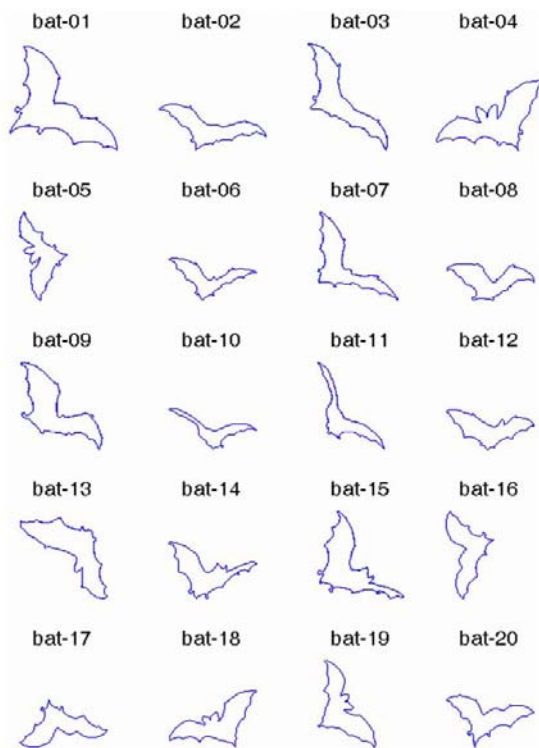


Fig. 17. Example of contours from the MPEG-7 Core Experiment CE-Shape-1. The changes between different patterns are due to the motion of the animal.

Table 3

Comparative results between IPM-based features and CSS-based ones in pattern retrieval

Method	Pattern	Average value of minimum similarity score	Average value of maximum similarity score	Average number of retained patterns
IPM	Bats	36.66	100.00	3.90
	Bell	10.85	100.00	20.00
	Children	11.55	100.00	20.00
	Global	19.77	100.00	14.63
	averages			
CSS	Bats	50.93	100.00	4.60
	Bell	42.52	100.00	4.50
	Children	48.17	100.00	10.70
	Global	47.20	100.00	6.60
	average			

The average value of the minimal similarity score together with the average value of the number of retained patterns indicates the ability of the method to recognise a pattern that has undergone rigid or non-rigid deformations: the higher these parameters are and the better is the method. As can be seen in Table 3, the two parameters seem to behave in opposition one of the other. While the IPM-based method provides the higher global average number of retained patterns and lower average of minimal similarity score, the CSS-based method provides the lower global average number of retained patterns and higher global average value of the minimum similarity score. That thus requires a compromise which will be guided by the contours under study.

5. Conclusion and discussion

We have presented some preliminary results of a new and simple method for planar pattern description within the context of pattern recognition. This paper analysis also the performance of some attributes derived from the IPM function. According to these results it would seem reasonable to believe that this is a good candidate method within the context of (planar) pattern recognition based on analysis of the closed non-occluded contour of the pattern being studied. This analysis method can be accelerated if just the IPM points for the concerned set of σ values are generated rather than generating the "entire" IPM function. This is an advantage as opposed to the CSS method which has to construct the entire CSS map in order to generate the features.

Similarly to the CSS method, our approach has some interesting advantages:

- Curves do not have to be convex;
- The pattern can be identified regardless of its orientation (the method is rotation insensitive);
- The method is translation insensitive;
- The method is scale change insensitive for a large range of scales.

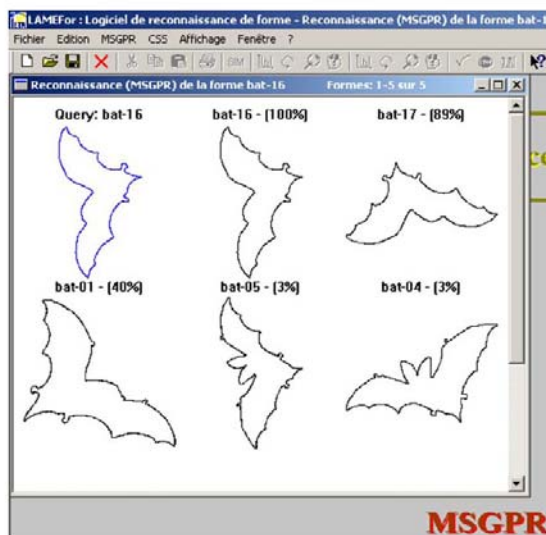


Fig. 18. Retrieval of a contour from the database of bats contours by using IPM-based features.

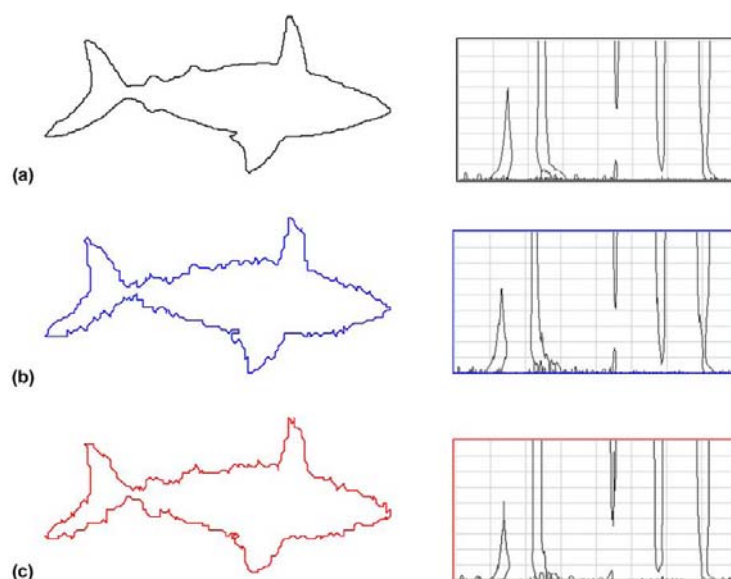


Fig. 19. The noise effect: (a) noise free, (b) additive noise with standard deviation $\sigma_n = 1.2$, (c) additive noise with standard deviation $\sigma_n = 1.8$.

The initial results obtained by applying this method to pattern retrieving from a database show that the IPM function produces consistent attributes and can therefore be positively applied to pattern recognition, especially within the context of OCR (or optical character recognition).

It is also interesting to note that the IPM function is very easy to understand and does not require much theory: this makes it very easy to implement. As stated above, our approach focuses on recognition of planar patterns that do not produce self-intersected contours. This aspect is therefore not taken into account in our studies.

Applying this method to noisy contours show that it is resistant to noise for a range of noise energy. From Fig. 19 we can note the following points: the IPM function is not too affected by noise for high scales (see upper part of the IPM functions); the apparition of many other structures in the lower part of the IPM function indicates that the pattern being studied is noisy. As stated in Section 2.5, high frequency information is located in the lower part of the IPM function: as the added noise is represented by fine details (with high frequency), this explains the appearance of those structures for low scales.

As stated earlier, the main weakness of our approach is that it fails to handle occluded contours and those having undergone a non-rigid deformation, nevertheless it is advantageous in that it is simple and easy to implement and to apply. It can also be considered as a useful step in a pattern recognition system, especially when one is concerned with rigid deformations.

Thanks to the results of the comparative study carried out between our method and an adapted-CSS method, we think that it is reasonable to say that our method has

its own place in the context of planar shapes description and recognition.

Acknowledgements

We are grateful to Mr. Sadegh Abasi, Dr. Farzin Mokhtarian and Prof. Josef Kittler for making their SQUID system publicly available.

References

- Belongie, S., Malik, J., Puzicha, J., 2002. Shape matching and object recognition using shape contexts. *IEEE PAMI* 24 (24), 509–522.
- Bruckstein, A., Katzir, N., Lindenbaum, M., Porat, M., 1992. Similarity invariant signatures for partially occluded planar shapes. *IJCV* 7 (3), 271–285.
- Bruckstein, A.M., Rivlin, E., Weiss, I., 1996. Recognizing objects using scale space local invariants. In: *Proceedings of the 1996 International Conference on Pattern Recognition (ICPR'96)*, Vienna, Austria, 25–29 August, pp. 760–764.
- Grigorescu, C., Petkov, N., 2003. Distance Sets for Shape Filters and Shape Recognition. *IEEE Trans. Image Process.* 12 (9).
- Kadyrov A., Petrou, M., 2001. Object descriptors invariant to affine distortions. In: *Proceedings of the British Machine Vision Conference, BMVC'2001*, Manchester, UK.
- Kpalma, K., Ronsin, J., 2003. A multi-scale curve smoothing for generalised pattern recognition (MSGPR). In: *Seventh International Symposium on Signal Processing and its Applications (ISSPA)*, Paris, France, pp. 427–430.
- Latecki, L.J., Lakamper, R., Eckhardt, U., 2000. Shape descriptors for non-rigid shapes with a single closed contour. In: *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 424–429.
- Lindeberg, T., 1994. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Lindeberg, T., 1998. Principles for automatic scale selection. Technical report ISRN KTH NA/P-98/14-SE, Department of Numerical Anal-

- ysis and Computing Science, KTH (Royal Institute of Technology), S-100 44 Stockholm, Sweden.
- Matusiak S., Daoudi M., 1998. Planar closed contour representation by invariant under a general affine transformation. In: IEEE International Conference on System, Man and Cybernetics (IEEE-SMC'98), Hyatt Regency La Jolla, San Diego, California, USA, 11–14 October, pp. 3251–3256.
- Mokhtarian, F., Bober, M., 2003. Curvature Scale Space Representation: Theory, Applications and MPEG-7 Standardization. Kluwer Academic.
- Mokhtarian, F., Mackworth, A.K., 1992. A theory of multiscale, curvature-based shape representation for planar curves. IEEE Trans. Pattern Anal. Machine Intell. PAMI-14 (8).
- Mokhtarian, F., Abasi, S., Kittler, J., 1996. Efficient and robust retrieval by shape content through curvature scale space. In: Proceedings of International Workshop on Image Databases and MultiMedia Search, Amsterdam, The Netherlands, pp. 35–42.
- Roh, K.-S., Kweon, I.-S., 1998. 2-D object recognition using invariant contour descriptor and projective refinement. Pattern Recognition 31 (4), 441–455.
- Sossa, H., 2000. Object Recognition, Summer School on Image and Robotics, INRIA Rhône-Alpes, France.
- Torres-Méndez, L.A., Ruiz-Suárez, J.C., Sucar, L.E., Gómez, G., 2000. Translation, rotation, and scale-invariant object recognition. IEEE Trans. Systems Man Cybernet. C 30 (1), 125–130.
- Trimeche, M., Alaya Cheikh, F., Gabbouj, M., 2000. Similarity retrieval of occluded shapes using wavelet-based shape feature. In: Proc. SPIE Internat. Symp. on Internet Multimedia Management Systems (VV10), Boston, MA, USA.
- Wang, Y.-P., Lee, S.L., Toraichi, K., 1999. Multiscale curvature-based shape representation using B-spline wavelets. IEEE Trans. Image Process. 8 (11), 1586–1592.
- Zhang, D., Lu, G., 2004. Review of shape representation and description techniques. Pattern Recognition 37, 1–19.

XX - Description de formes par analyse multiéchelle de l'angle tangentiel

Kidiyo Kpalma and Joseph Ronsin

Turning Angle Based Representation for Planar Objects

Electronics Letters

10 May 2007

Vol. 43, No 10, pp. 561-563 (4 pages)

Résumé

Nous présentons une nouvelle approche pour la reconnaissance des formes. Basée sur l'analyse multi-échelle des contours fermés, cette méthode utilise l'angle tangentiel. A l'aide d'un lissage progressif, du contour d'entrée, on génère la carte de l'angle tangentiel différentiel ou differential-Turning Angle Scale Space (d-TASS) pour décrire ce contour. Des résultats expérimentaux montrent que la carte d-TASS est fortement caractéristique du contour et qu'elle est résistante aux transformations affines. La carte d-TASS est donc bien adaptée pour la description des formes.

Turning Angle Based Representation for Planar Objects

KPALMA K. and J. RONSIN

We present a new approach for pattern description. Based on a multi-scale analysis of closed contours, this method deals with the differential turning angle extracted from progressively smoothed contour to generate the differential-Turning Angle Scale Space (d-TASS) map. Experimental results show that the d-TASS map is closely related to the contour and it is resistant to affine transformation so it is well-suited for pattern description.

Introduction: An automatic pattern recognition system must be simple, fast and efficient since one deals, in general, with large databases. The method must be resistant to affine transformations and noise. Moreover, it must also be robustly resistant to occultation [3, 5]. In this paper, we present an original representation called the d-TASS function (differential-Turning Angle Scale Space) that helps to assess the curvature [4] which can be expressed in terms of turning angle [6, 7]. This CSS-like [1, 4] method is also based on the multi-scale smoothing of the boundary of a planar pattern. It is translation insensitive, rotation invariant, scale change invariant and shearing resistant.. Reference [7] describes a method based on the (multiscale) turning angle in which features are defined by the turning angles of two different scales. In the new approach, we propose some essential characteristic points in the d-TASS map that yield representative features.

Description of the approach: The input contour is considered to be defined by a set of N points ordered counterclockwise in the plane. Given a starting

point P_0 , the points are numbered from P_0 to P_{N-1} . Now, let \vec{V}_n be the vector defined by the two points P_n and P_{n+1} originating at P_n and oriented towards P_{n+1} . Figure 1 illustrates an example of this. The turning angle θ_n at point P_n is the angle between vectors \vec{V}_n and the x-axis. The differential turning angle (d-TA) is the angle φ_n between the two consecutive vectors \vec{V}_n and \vec{V}_{n-1} :

$$\varphi_n = \theta_n - \theta_{n-1} \quad (1)$$

The d-TA is positive at a convex point, negative at concave points and null on straight line. The locations of the maxima and the minima of d-TA correspond to points with high curvature magnitude, constituting the essential points of the contour. Figure 2 illustrates some properties of d-TA: the high value of d-TA in (A) indicates a convex portion (a) of the contour. Part (c), which is a convex arc, gives a positive d-TA (C) while part (i), which is a concave arc, gives a negative d-TA (I). Straight sections (f) give null d-TA (F).

Definition: We define the essential points, which can be classified into two categories:

- α -points correspond to those with a local minimum d-TA (concavity),
- β -points correspond to those with a local maximum d-TA (convexity).

J. Feldman et al. demonstrated in [2] that the sign of the curvature carries useful information. The d-TA function therefore provides a particularly suitable means of retrieving useful features along the contour.

In [4], the authors showed that scale space analysis can be used to immediately capture useful information about multiple scales. In order to take advantage of this property, we derive a multiscale version of the d-TA to define the differential turning angle scale space (d-TASS) function.

Let the d-TASS function be a 2-D function giving the differential turning angle. The horizontal axis corresponds to the space variable u while the vertical axis is related to the scale variable σ . The **d-TASS map** is then defined by the essential points (see Definition) detected from the d-TASS function.

This map indicates the location of the essential points on a smoothed contour for a filter standard deviation σ . The principle of this approach can be described as follows:

1. the input contour is separated into its parameter-based coordinate functions $x(u)$ and $y(u)$,
2. both functions are smoothed by applying a Gaussian low-pass filter with standard deviation σ ,
3. the d-TASS map is then generated by detecting the location of the essential points from the d-TASS function.

Figure 3.a shows the d-TASS function for a range of the filter standard deviation σ . At each filtering iteration, the d-TA along the filtered contour is normalized to obtain gray levels between 0 and 255. As shown in this figure, the resulting image yields some particularly interesting observations: on this image, dark areas indicate the location of concave sections and clear areas indicate convex sections (see Fig.3.b). **The d-TASS map yields interesting properties such as:**

Shape representativity: Figure 4 shows the d-TASS maps derived from various shapes. It is evident that representing the turning angle in the scale space is characteristic of the contour.

Rotation insensitivity: In figure 5 we represent the d-TASS maps derived from the same contour after subjecting it to rotation. It is clear that the d-TASS map is topologically rotation insensitive: the rotation is just expressed by a simple circular shift.

Resistance to scale change: Figures 6 show examples of d-TASS maps extracted from a contour that has undergone a scale change: these maps remain practically the same whatever the scale factor.

Resistance to occultation: On figure 7 we represent the d-TASS map of a contour (Fig.7.a) and the d-TASS map derived from the same contour after a partial occultation(Fig.7.b). As is obvious from these figures, the d-TASS map from the remaining visible part is the same as the part derived from the corresponding region on the whole contour. In these figures, RO indicate the same structures from d-TASS mapping of the remaining visible part and the corresponding ones in the d-TASS map derived from the whole contour. This property is particularly important and interesting since it makes it possible to recognize a pattern even if it is partially hidden.

Shearing resistance: Figure 8 shows examples of the effect of shearing on the d-TASS map. In this study, the shearing transformation is performed as follows:

$$\begin{pmatrix} x_s \\ y_s \end{pmatrix} = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_0 + \lambda y_0 \\ y_0 \end{pmatrix} \quad (2)$$

where (x_0, y_0) are the coordinates of the original contour and (x_s, y_s) those of the sheared equivalent. These figures indicate that the d-TASS map derived from the sheared contours is consistent with the map extracted from the original contours. Except for slight loss in the height of some peaks, the d-TASS maps are topologically identical.

Application of the d-TASS map: From the results obtained from various experiments above, we observe that the d-TASS map is closely related to the contour: it is characteristic of the shape under consideration. In addition, it is

resistant to affine transformations. We have also showed that for partially occluded contours, the visible region generates a d-TASS map that remains consistent with the area corresponding to the same region on the d-TASS map derived from the whole contour. With all these interesting properties one can define efficient features based on the d-TASS map to describe an object. Possible features are the horizontal distances between two essential points at a given scale [1] or the location (u_i, σ_i) of points where essential points merge.

Conclusion: It is well established that curvature carries useful information about contours. In addition, it is proven that the capacity of a system to simultaneously handle various scales makes it easy to capture useful features from an object. Based on this, the d-TASS map appears to be suitable for planar shape description.

The originality of this approach is its ability to deal with multiple scales at the same time by utilizing the differential turning angle in the scale space representation.

References

- [1] Kpalma K. and J. Ronsin, 'Multiscale contour description for pattern recognition', Elsevier, Pattern Recognition Letters, Vol.27, No.13, pp. 1545-1559, 1 October 2006
- [2] Feldman J. and M. Singh, 'THEORETICAL NOTE: Information Along Contours and Object Boundaries', Psychological Review, The American Psychological Association, 2005, Vol. 112, No. 1, pp. 243-252
- [3] Bruckstein, A. M., Rivlin, E., and Weiss, I., 1996. 'Recognizing objects using scale space local invariants', Proceedings of the 1996 International Conference on Pattern Recognition (ICPR '96), August 25-29, pp. 760-764, Vienna, Austria.
- [4] Lindeberg T., 'Scale-Space Theory in Computer Vision', Kluwer Academic Publishers, Dordrecht, Netherlands, 1994.
- [5] Bruckstein A., N. Katzir, M. Lindenbaum and M. Porat, 'Similarity-Invariant Signatures for Partially Occluded Planar Shapes', IJCV, Vol. 7/3, pp. 271-285, 1992.
- [6] Niblack W. and J. Yin, 'A pseudo-distance measure for 2D shapes based on turning angle', Proceedings of the International Conference on Image Processing (ICIP'95), Vol.3, pp. 352-355, 23-26/10/1995.
- [7] Esther M. Arkin, L. Paul Chew, Daniel P. Huttenlocker, Klara Kedem, and

Joseph S. B. Mitchell, 'An efficiently computable metric for comparing polygonal shapes', IEEE Transactions on PAMI, 13(3), pp. 209–216, March 1991.

Authors' affiliations:

KPALMA K. and J. RONSIN

IETR (Institut d'Électronique et de Télécommunications de Rennes)

UMR – CNRS 6164

Groupe Image et Télédétection

Institut National des Sciences Appliquées (INSA) de Rennes

kidiyo.kpalma@insa-rennes.fr

Figure captions

Fig.1: Definition of a turning angle

Fig.2: Convexity/concavity on a planar object and corresponding d-TA

Fig.3: The d-TA function vs scale-space

Fig.4: The d-TASS map of different contours

Fig.5: The d-TASS map vs the rotation

Fig.6: The d-TASS map vs the scaling

Fig.7: The d-TASS map vs occultation

Fig.8: The d-TASS map vs shearing

Figure 1

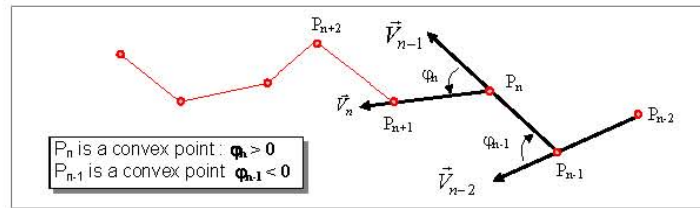


Figure 2

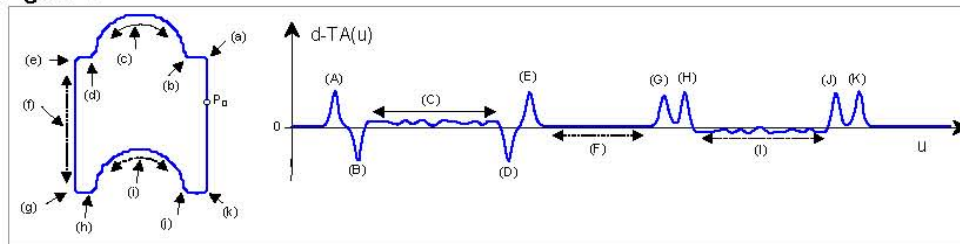
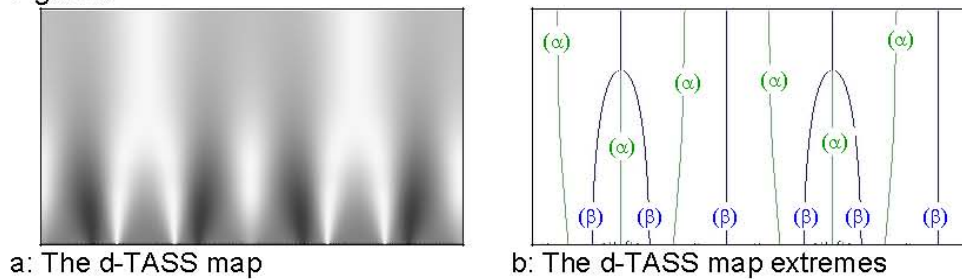


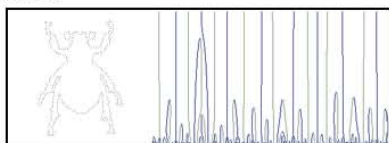
Figure 3



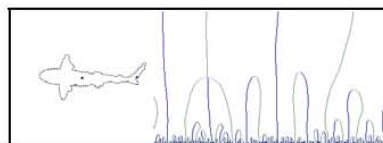
a: The d-TASS map

b: The d-TASS map extremes

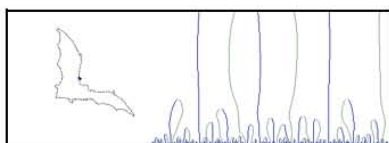
Figure 4



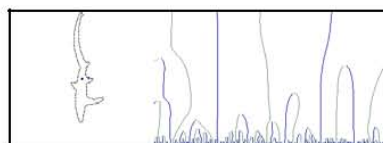
a: beetle-05



b: kk20

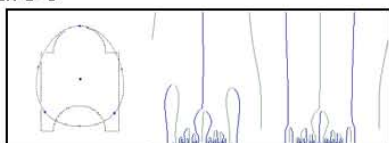


d: bat-07

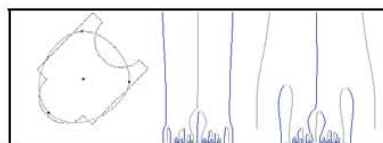


d: kk711

Figure 5

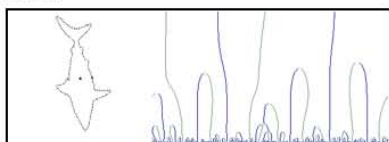


a: 0° rotation

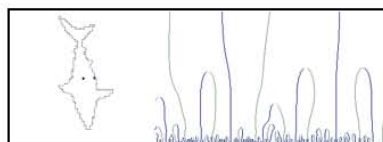


b: 135° rotation

Figure 6

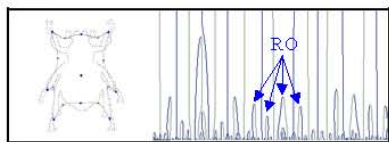


a: original

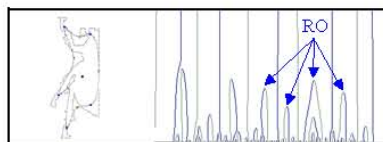


b: scaled (x0.20)

Figure 7

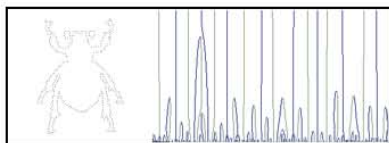


a: Original

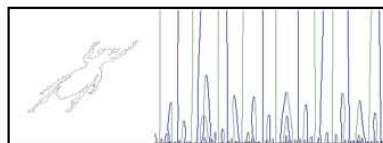


b: Right side occultation

Figure 8



a: Original



b: Right side occultation

Habilitation à diriger des recherches

Kidiyo KPALMA

Titre : Analyse et Interprétation d'Images : Analyse Multi-échelle pour la Description des Formes Planaires.

Résumé : L'essor technologique, induit une production d'énormes masses de données électroniques et oblige à une adaptation des traitements du flux d'information correspondant. D'où un besoin d'outils d'aide à l'interprétation permettant des prises de décision dans les délais raisonnables. En matière d'imagerie numérique, les champs d'application sont vastes allant de l'utilisation domestique (appareil photo numérique, webcam ou simplement téléphone portable) jusqu'aux applications de type télédétection. L'évolution croissante du volume des données s'accompagne de contraintes, qu'il s'agisse du dépouillement des images de télédétection et c'est là que se situe un premier axe de nos travaux, ou du tri de photos, ou de la recherche d'une photo spécifique, prise à une certaine occasion en un lieu donné et c'est là le second axe de nos travaux.

S'agissant de la télédétection, la multitude de satellites d'observation nous fournit une grande variété d'images. Celles-ci nous apportent des informations à la fois redondantes et complémentaires. Aussi s'avère-t-il judicieux de les combiner dans un processus de fusion pour en tirer des renseignements plus consistants et plus intelligibles.

Traditionnellement, l'indexation et la recherche d'une photo se fait à l'aide d'une annotation par mots-clés. Compte-tenu des limitations de cette procédure, l'enjeu, aujourd'hui, est la détermination de descripteurs efficaces basés sur le contenu des images pour leur indexation.

Toutes ces considérations visent à proposer des solutions alliant différents descripteurs dans un processus de fusion pouvant fournir un index hybride prenant en compte l'aspect sémantique du contenu de l'image. C'est dans ce contexte d'analyse d'images pour leur interprétation que se situent nos travaux de recherche dont le présent mémoire fait l'objet. Il s'articule en trois parties. Dans la première partie, nous présentons un résumé des activités tant du point de vue de l'enseignement que de la recherche : un aperçu de nos travaux de recherche retrace les études réalisées pour des applications de fusion en télédétection et pour la description des formes. La deuxième partie reprend plus en détails l'une de nos études sur le thème de l'analyse multi-échelle de contours d'objets planaires pour des applications de description et de reconnaissance de formes. On y présente ainsi la méthode MSGPR basée sur la carte des points d'intersection, méthode que nous proposons pour l'extraction d'un descripteur pour l'indexation des objets planaires. Pour finir, la dernière partie présente quelques articles représentatifs de nos activités de recherche : analyse de textures, recalage d'images, fusion d'image, analyse espace-échelle, reconnaissance de formes.

Mots-clé : analyse et interprétation, analyse multi-échelle, scalogramme, reconnaissance de formes, indexation, analyse de textures, segmentation d'image, fusion d'images, télédétection